



SISTEMAS DE INFORMAÇÃO E CARTAS DE CONTROLO PARA MONITORIZAÇÃO DE UM PROCESSO PRODUTIVO

por

Joana Cristina Costa Carvalho

Relatório de Mestrado em Modelação, Análise de Dados e Sistemas de
Apoio à Decisão

Orientada por

Professora Doutora Fernanda Otilia Figueiredo

Faculdade de Economia

Universidade do Porto

2015

Dedicado a Rita Assunção e aos meus pais!

Nota Biográfica

Joana Cristina Costa Carvalho ingressou no Ensino Superior no ano de 2009, na Licenciatura em Matemática que finalizou com Formação Complementar (*minor*) em Matemática Aplicada, na Faculdade de Ciências da Universidade do Porto - FCUP.

Uma vez que se identificava com a vertente empresarial optou por ingressar no Mestrado de Modelação, Análise de Dados e Sistemas de Apoio à Decisão, na Faculdade de Economia da Universidade do Porto - FEP, em Setembro de 2013.

Para finalizar o mestrado concretizou o estágio curricular na maior empresa portuguesa do setor da cortiça, Corticeira Amorim S.G.P.S., S.A.. Este foi realizado na Amorim & Irmãos, S.A., Unidade de Negócios de Rolhas.

Agradecimentos

O meu sincero e profundo agradecimento pela contribuição nesta etapa fundamental da minha formação:

À Amorim & Irmãos, S.A. pela oportunidade de realização do estágio curricular, por todos os recursos e oportunidades disponibilizadas.

Particularmente à Engenheira Cristina Cardoso por todo o apoio, disponibilidade e integração ao longo desta etapa.

À equipa do laboratório da Amorim Distribuição pela disponibilidade, ajuda e boa disposição.

Ao Senhor Rodrigo Cordeiro pela paciência e partilha de experiências importantes para este trabalho.

Desde os diretores de qualidade aos empregados fabris agradeço os ensinamentos e por permitirem vivenciar o que é pertencer à grande equipa que é a Amorim & Irmãos, S.A..

À Professora Doutora Fernanda Figueiredo por toda a partilha de conhecimento, auxílio prestado no acompanhamento e enriquecimento deste projeto.

Aos meus pais pelo incentivo, diretrizes e apoio incondicional, em todos os momentos ao longo desta jornada.

À minha irmã pelas horas que passamos no debate de ideias e amparo nas várias fases.

Ao Diogo Marques pelo incentivo, dedicação, confiança que sustentaram a realização deste trabalho.

À minha família pela preocupação e lições demonstradas nos diversos momentos da minha formação.

Aos meus amigos que auxiliaram, direta ou indiretamente, e deixaram marca no meu percurso académico. De forma particular Ana Catarina Ramos, Ana Margarida Silva, Ana Rita Teixeira, Filipa Freire, Maria Idalina Santos e Joana Gomes, o meu grande obrigado.

Tabela de Símbolos Matemáticos

Notação Matemática	Descrição
R	Amplitude amostral
λ	Constante de Amortecimento
k	Dimensão do subgrupo
n	Dimensão da amostra
σ	Desvio-padrão populacional
s	Desvio-padrão amostral
α	Erro de tipo I ou nível de significância
β	Erro de tipo II
W_t	Estatística de controle
Z_i	Estatística EWMA
L	Fator de multiplicação de desvio-padrão
H_0	Hipótese nula
H_1	Hipótese alternativa
np	Número de não conformes
\bar{R}	Média das amplitudes dos subgrupos
\bar{S}	Média dos desvio-padrão dos subgrupos
$\bar{\bar{X}}$	Média dos valores individuais do subgrupo
$\bar{\bar{X}}$	Média das médias dos subgrupos
μ	Média populacional
σ_0	Valor alvo para o desvio-padrão do processo produtivo
μ_0	Valor alvo para a média do processo produtivo
T	Valor nominal da especificação

Abreviaturas

AD	Amorim & Irmãos, S.A. - Unidade Industrial Amorim Distribuição
AI	Amorim & Irmãos, S.A. - Unidade Industrial Lamas
ARL	<i>Average Run Length</i>
BI	<i>Business Intelligence System</i>
CEP	Controlo estatístico do processo
CIPR	Código Internacional das Práticas Rolheiras
CK	Amorim & Irmãos, S.A. - Unidade Industrial Champcork
CUSUM	<i>Cumulative Sum</i>
DAS	<i>Data Area Stating</i> (Área de Estacionamento de dados)
DM	<i>Data Marts</i>
DS	Amorim & Irmãos, S.A. - Unidade Industrial De Sousa
DW	<i>Data Warehouse</i>
EQ	Amorim & Irmãos, S.A. - Unidade Industrial Equipar
ERP	<i>Enterprise Resource Planning</i>
ETL	<i>Extracting, Transforming and Loading</i>
EWMA	<i>Exponentially Weighted Moving Average</i>
FSI	<i>Fixed Sampling Interval</i>
IT	Intervalo de Tolerância
LC	Linha central
LCI	Limite de Controlo Inferior
LCS	Limite de Controlo Superior
LEI	Limite de Especificação Inferior
LES	Limite de Especificação Superior
LVI	Limite de Vigilância Inferior
LVS	Limite de Vigilância Superior
MA	<i>Moving Average</i> (Média Móvel)
ODBC	<i>Open Database Connectivity</i>
OF	Ordem de Fabrico
OLAP	<i>Online Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>

PC	Posto de Controlo
PK	Amorim & Irmãos, S.A. - Unidade Industrial Portocork
QUALIS	<i>Quality Information System</i>
SI	Sistema de Informação
SPC	<i>Statistical Process Control</i>
TCA	2,4,6 - Tricloroanisol
TS	Amorim & Irmãos, S.A. - Unidade Industrial TopSeries Unit
UI	Unidade industrial
VBA	<i>Visual Basic for Applications</i>
VL	Amorim & Irmãos, S.A. - Unidade Industrial Vasconcelos & Lincke
VSI	<i>Variable Sampling Interval</i>

Resumo

Num mercado competitivo, as empresas necessitam ter um sistema de informação coeso dos seus produtos e processos produtivos, de forma fácil, rápida e dinâmica, para serem capazes de ultrapassar as várias pressões a que são sujeitas. Desta forma, cada vez mais procuram ter ao seu dispor os sistemas de informação mais adequados, para os seus utilizadores adquirirem todas as informações necessárias às diversas tomadas de decisão.

Dois sistemas de informação utilizados na empresa Amorim & Irmãos, S.A., no departamento de Qualidade, serão abordados neste projeto. Também serão apresentadas algumas validações de procedimentos já existentes e lacunas detetadas que foram retificadas e/ou encontram-se em fase de investigação ou implementação, com o intuito de possuir sistemas de informação mais incisivos na informação transmitida.

Quando se trata da qualidade de um produto, não pode ser apenas considerado o produto finalizado. Para a obtenção de um produto de excelência todas as etapas, desde a matéria-prima ao produto acabado, têm de ser vigiadas, de modo que o Controlo Estatístico do Processo não pode ser descurado. Com o controlo estatístico é possível a partir de uma amostragem inferir qual é o cenário da população. Neste âmbito é apresentado um caso de estudo real, num ponto fulcral para a Qualidade de um processo produtivo da empresa. Para isso foram utilizadas as cartas de controlo (Shewhart e EWMA) para a monitorização e controlo do processo, por forma a responder corretamente às especificações impostas.

Palavras-Chave: Sistemas de Informação, Processamento Analítico Online, Carta de Controlo de Shewhart, Carta de Controlo EWMA

Abstract

In a competitive market, the companies need to have a cohesive information system of their products and production process in an easy, fast and dynamic manner, so that they can overcome the many pressures that are subjected. Thus, increases the search to have in their disposal the most appropriate information systems for their users to acquire all the necessary information for the various decision-making.

Two information systems used in the company Amorim & Irmãos, S.A., in the Quality department, were introduced in this project. Also some validations of existing procedures and detected gaps were presented that were afterwards rectified and/or are in stages of research or implementation, in order to have more accurate information systems for transmitted information.

As far as the quality of a product is concerned, it can not only, be considered the finished product but also, to obtain an excellent product every stage, from raw material to finished product, supervision is demanded, so that the Statistical Process Control can not be overlooked. With statistical control it is possible from a sample to infer which is the behavior of the population. In this context it was presented a real case study, focusing an important point for the quality of a production process of the company. For this, control charts (Shewhart and EWMA) were use for monitoring the process control in order to answer correctly the specifications imposed.

Keywords: Information systems, *Online Analytical Processing*, Shewhart Control Chart, EWMA Control Chart

Índice

Nota Biográfica	ii
Agradecimentos	iii
Tabela de Símbolos Matemáticos	iv
Abreviaturas	v
Resumo	vii
Abstract	viii
1 Introdução	1
1.1 Motivação	1
1.2 Apresentação da Empresa Corticeira Amorim S.G.P.S., S.A.	2
1.2.1 Sistemas de informação	4
1.2.2 Processo produtivo	4
1.3 Organização do relatório	5
2 Estado da Arte	6
2.1 Sistemas de Informação	6
2.1.1 <i>Online Transaction Processing</i>	7
2.1.2 <i>Online Analytical Processing</i>	7
2.1.2.1 <i>Data Warehouse</i>	8
2.1.2.2 <i>Data Marts</i>	8
2.1.2.3 Extração, Transformação e Transporte	9
2.1.2.4 Modelação Dimensional	10
2.1.2.5 Metadados	12
2.2 Cartas de Controlo	13
2.2.1 Cartas de Controlo de Shewhart	17
2.2.1.1 Coeficientes de assimetria e achatamento	17
2.2.1.2 Testes de Normalidade	18
2.2.2 Cartas de Controlo EWMA	24

2.2.3	Capacidade de Processo	26
2.2.3.1	Capacidade de Processo para distribuições Normais .	27
2.2.3.2	Capacidade de Processo para distribuições Não Nor- mais	29
3	Trabalho desenvolvido no estágio curricular	34
3.1	Sistemas de informação	34
3.1.1	Uniformização da informação	35
3.1.2	Acompanhamento a fornecedores	38
3.1.3	Quantificação de TCA	41
3.1.4	Caraterística Volátil do QUALIS	43
3.1.5	Alteração de ensaios <i>ControLab</i>	44
3.1.6	Alteração ao <i>SPC</i>	47
3.1.7	Registos do <i>SPC</i> no QUALIS	53
3.2	Controlo Estatístico do Processo - Caso de Estudo	54
3.2.1	Amostragem	55
3.2.2	Cartas de Controlo de Shewhart	55
3.2.2.1	Estudo amostragens no ROSA EVO	58
3.2.2.2	Transformação Box-Cox	71
3.2.2.3	Índices de Capacidade	76
3.2.2.4	Análise de Correlação	79
3.2.3	Cartas de Controlo de EWMA	83
4	Conclusões e Trabalho Futuro	87
4.1	Conclusões	87
4.2	Trabalho Futuro	88
	Bibliografia	90

Lista de Tabelas

2.1	Operações dimensionais em DW	11
2.2	Causas de variabilidade no CEP	13
2.3	Valores referência para índice C_p	27
3.1	Quantidade de entidades presentes no QUALIS, em função da UI . .	37
3.2	Índices de Capacidade para valores individuais e médios dos subgrupos	49
3.3	Média, Desvio-padrão, coeficiente de assimetria e coeficiente de achatamento, dos registos de inspeção de 2014	57
3.4	Medidas de dispersão, dos registos de inspeção de 2014	57
3.5	Resultado do teste Kolmogorov-Smirnov dos registos de inspeção de 2014	58
3.6	Distribuição lotes pela Estufas ROSA EVO	62
3.7	Média, Desvio-padrão, coeficiente de assimetria e Coeficiente de achatamento, dos registos de controlo em Janeiro de 2015	64
3.8	Medidas de dispersão, dos registos de controlo em Janeiro de 2015 . .	65
3.9	Resultado teste de Normalidade de Kolmogorov-Smirnov, Janeiro 2015	66
3.10	Limites de Controlo e np , dados originais de Janeiro 2015	68
3.11	Medidas de dispersão, das médias amostrais de Janeiro de 2015 . . .	69
3.12	Média, Desvio-padrão, coeficiente de assimetria e Coeficiente de achatamento, das médias amostrais de Janeiro de 2015	69
3.13	Resultados do teste de Kolmogorv-Smirnov e Shapiro-Wilk, aplicado às médias dos subgrupos de Janeiro 2015	70
3.14	Valores de λ para transformação dos dados	71
3.15	Média, Desvio-padrão, coeficiente de assimetria e coeficiente de achatamento após Transformação Box-Cox, Janeiro de 2015	71
3.16	Medidas de dispersão, dos registos de inspeção de Janeiro de 2015 transformados	72
3.17	Resultados teste de Kolmogorov-Smirnov, dados transformados Janeiro 2015	73
3.18	Limites de Controlo e número amostragens fora de controlo, Janeiro 2015	76
3.19	Desvio-padrão associado às estufas ROSA EVO, Janeiro 2015	77

3.20	Índices de Capacidade do Processo para distribuições Normais, dados originais Janeiro 2015	77
3.21	Percentis associados às estufas ROSA EVO, Janeiro 2015	78
3.22	Índices de Capacidade do Processo para distribuições Não Normais, Janeiro 2015	78
3.23	Desvio-padrão dados transformados, Janeiro 2015	79
3.24	Índices de Capacidade do Processo para distribuições Normais, dados transformados Janeiro 2015	79
3.25	Coeficiente Spearman entre Média amostral antes e após ROSA EVO	82
3.26	Coeficiente Spearman entre desvio-padrão amostral antes e após ROSA EVO	82
3.27	Não Conformidades carta de controlo \bar{X} e EWMA	85
3.28	Não Conformidades para valores de referência de λ e L , na carta de controlo EWMA de Janeiro de 2015	85

Lista de Figuras

1.1	Unidades de Negócio da Corticeira Amorim S.G.P.S, S.A.	2
1.2	Tipos de rolhas produzidas na Amorim & Irmãos, S.A.	3
1.3	Processo Produtivo de Rolha Natural	4
1.4	Processo Produtivo de Rolha Aglomerada com discos	5
2.1	<i>Data Warehouse</i> com <i>Data Marts</i>	8
2.2	Ilustração de um <i>Data Warehouse</i> com área de estacionamento	9
2.3	Ilustração de esquemas Estrela e <i>Snowflake</i>	10
2.4	Ilustração de operações dimensionais	12
2.5	Exemplo de carta de controlo	15
2.6	Processo estável versus Processo instável	16
2.7	Ilustração de coeficientes de assimetria e achatamento	18
2.8	Função Perda de Taguchi	29
3.1	Ficha entidade do <i>Controlab</i>	36
3.2	Impacto da Consolidação das entidades presentes no QUALIS	38
3.3	Análise Mensal / Fornecedores	39
3.4	Análise Anual dos Principais Fornecedores	40
3.5	Alerta introduzido no <i>SPC</i> para subgrupos de dimensões diferentes .	47
3.6	Visualizações dos cálculos efetuados com valores individuais no <i>SPC</i> .	48
3.7	Listagem dos subgrupos Carta de Controlo	48
3.8	Carta de Controlo $\bar{X} - R$ e respetivos índices de capacidade, ROSA EVO 3	49
3.9	Índices de capacidade mais usuais	50
3.10	Histograma Valores individuais	51
3.11	Especificações para os valores individuais	51
3.12	Carta de Controlo $\bar{X} - R$	52
3.13	Carta de Controlo $\bar{X} - S$	53
3.14	Histogramas de dados Primevara/Verão de 2014	56
3.15	Histogramas de dados Outono/Inverno de 2014	56
3.16	Amostragem Superficial após Lavação	59
3.17	Amostragem Superficial VS Profundidade após Lavação	60
3.18	Amostragem Atual VS Proposta após Lavação	61

3.19 Média das amostragens por lote	61
3.20 Diferença de humidade entre fundo e cimo contentor, antes estufa . .	62
3.21 Amostragens distribuição pelo tapete	63
3.22 Amostragem Atual VS Proposta após Estufa	64
3.23 Histogramas dos dados de Janeiro 2015	65
3.24 Carta de Controlo de Shewhart para estufa 1	66
3.25 Carta de Controlo de Shewhart para estufa 2	67
3.26 Carta de Controlo de Shewhart para estufa 3	67
3.27 Carta de Controlo de Shewhart para estufa 4	68
3.28 Distribuição das médias dos subgrupos de Janeiro 2015, das diversas Estufas ROSA EVO	70
3.29 Valores individuais dos dados transformados de Janeiro 2015	72
3.30 Ajustamento à distribuição Normal dos dados transformados de Ja- neiro 2015	73
3.31 Carta de Controlo de Shewhart para estufa 1	74
3.32 Carta de Controlo de Shewhart para estufa 2	74
3.33 Carta de Controlo de Shewhart para estufa 3	75
3.34 Carta de Controlo de Shewhart para estufa 4	75
3.35 Distribuição dos valores médios das amostragens	80
3.36 Teste Normalidade para média amostrais de humidade	81
3.37 Gráfico de dispersão da média amostral antes <i>versus</i> após ROSA EVO	81
3.38 Gráfico de dispersão do desvio-padrão amostral antes <i>versus</i> após ROSA EVO	82
3.39 Cartas de Controlo EWMA para $\lambda = 0.25$ e $L = 2.998$	84
3.40 Cartas de Controlo EWMA para $\lambda = 0.05$ e $L = 2.615$	86

Capítulo 1

Introdução

1.1 Motivação

O estágio curricular foi realizado na Corticeira Amorim, S.G.P.S., S.A., na Unidade de Negócio Rolhas - Amorim & Irmãos, S.A., no departamento da Qualidade. Teve como objetivo a melhoria e desenvolvimento dos sistemas de informação no controlo do produto e cartas de controlo para monitorização dos processos produtivos.

Atualmente as empresas precisam de acompanhar pormenorizadamente os resultados dos seus processos produtivos, uma vez que para obterem vantagem competitiva sob os seus concorrentes é fulcral que tenham acesso aos seus dados detalhados e sobre várias perspetivas. Isto é, as empresas para serem bem sucedidas necessitam de aliar produtos de qualidade com o conhecimento histórico dos processos produtivos, sendo para isso fundamental terem sistemas de informação eficazes que as apoiem nas tomadas de decisão.

O sistema de informação QUALIS (*Quality Information System*) é um pilar que auxilia a empresa no âmbito da qualidade, pois este reúne a informação dos registos efetuados nos *softwares ControLab* e *SPC (Statistical Quality Control)*, como será explicado no decorrer deste trabalho. A informação proveniente do *ControLab* já está mais uniformizada de forma a poder ser analisada transversalmente para as várias unidades industriais (UI's) da Amorim & Irmãos, S.A.. Porém, cada UI possui especialização para diferentes produtos para os quais é necessário a realização de vários ensaios laboratoriais específicos. Assim existem sempre algumas diferenças nas necessidades de registo e tratamento de informação para cada tipo de rolha em cada UI. Consequentemente, ainda há pontos a melhorar, tais como:

- Consolidar as informações para todas as UI's e identificação de alteração dos registos para o mesmo item (designações dos fornecedores, produtos, categoria de relatórios) para existir coerência entre estas, sem eliminar as especificidades de cada UI;

- Criação de *reports* adequados às necessidades de cada UI, para eliminar utilizações de ODBC (*Open Database Connectivity*) sobre os *softwares* de qualidade atualmente em uso.

Tal como já foi referido anteriormente, o QUALIS retira informação proveniente do *SPC*, *software* utilizado para o controlo estatístico do processo. Contudo, grande parte da informação não está visível de forma adequada e/ou completa. Assim, para o bom desempenho do sistema de informação devem ser analisados quais os campos extraídos incorretamente e os que são ainda necessários extrair para o QUALIS.

Em *SPC* as cartas de controlo são uma ferramenta muito poderosa no controlo de um processo produtivo, pois permitem detetar alterações no processo ao longo do tempo. A identificação destas alterações permitirá tomar ações para reduzir a variabilidade das características do produto.

Outro objetivo do estágio curricular foi efetuar um estudo das cartas de controlo de Shewhart, utilizadas atualmente na empresa, e EWMA (*Exponentially Weighted Moving Average*) para utilização na monitorização dos processos de produção, de forma a avaliar qual será a mais adequada nos seus processos.

1.2 Apresentação da Empresa Corticeira Amorim S.G.P.S., S.A.

A Corticeira Amorim é a maior empresa mundial de produtos de cortiça e encontra-se organizada em cinco Unidades de Negócios, como ilustrado na figura 1.1:

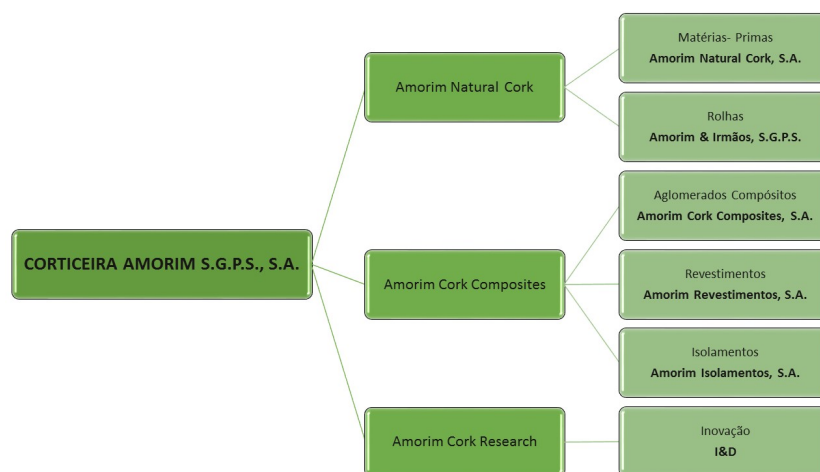


Figura 1.1: Unidades de Negócio da Corticeira Amorim S.G.P.S, S.A.

A Unidade de Negócio de rolhas tem produção média anual de quatro mil milhões de rolhas, correspondendo à unidade de negócios com mais vendas em 2013 o que prefaz 60,7% das vendas da Corticeira Amorim (Amorim, 2013).

A missão da empresa é “Acréscitar valor à cortiça, de forma competitiva, diferenciada e inovadora, em perfeita harmonia com a Natureza.” (Amorim, 2013).

A Amorim & Irmãos, S.A. tem um vasto portefólio de rolhas para responder a todos os segmentos de bebidas (vinhos, espirituosos e espumantes). Tem um segmento para bebidas espirituosas, espumantes e até mesmo uma gama de rolhas luxuosas. Todas estas rolhas têm em comum a sua matéria-prima, a cortiça.



Figura 1.2: Tipos de rolhas produzidas na Amorim & Irmãos, S.A.

A Unidade de Negócio de rolhas é constituída por 8 UI's, em que cada uma tem a sua especialização:

- Amorim & Irmãos - Lamas (AI): Produção de Rolhas Acquamark®, Colmatadas e Naturais;
- Amorim Distribuição (AD): Marcação e tratamento de Rolhas Acquamark®, Colmatadas, Helix®, Naturais, Neutrocork® e Twin Top®;
- Champcork (CK): Produção, marcação e tratamento de Rolhas de Champahe - Spark®;
- De Sousa (DS): Produção de Rolhas Helix®, Neutrocork®;
- Equipar (EQ): Produção, marcação e tratamento de Rolhas Advantec®, Aglomeradas e Twin Top®;
- Portocork (PK): Lavação, escolha, marcação e tratamento de Rolhas Naturais;
- TopSeries (TS): Produção de Rolhas capsuladas e especialidades;
- Vasconcelos & Lincke (VL): Lavação, escolha, marcação e tratamento de Rolhas Colmatadas e Naturais.

1.2.1 Sistemas de informação

No departamento da Qualidade, todas as UI's têm um laboratório, onde são realizadas várias análises de controlo aos produtos, ao longo do processo produtivo. A Amorim & Irmãos, S.A. tem certificações como o CIPR (Código Internacional das Práticas Rolheiras) e as ISO 9001 (Qualidade) e 22000 (Segurança Alimentar), que garantem a qualidade do produto final. As análises realizadas incidem sobre as dimensões e massa (das rolhas), humidade, oxidantes, forças de extração e quantificação do composto 2,4,6-Tricloroanisol (TCA), e outras.

Cada laboratório possui o *software ControLab* onde são registados os dados de todas análises realizadas às rolhas, permitindo o controlo do produto nas várias fases do processo produtivo, e elaboração de relatórios de qualidade para o cliente. Como referido anteriormente cada rolha tem a sua especificidade, dada esta condição o *ControLab* apesar de ser semelhante em todas as UI's a arquitetura das bases de dados são diferentes.

O elevado número de dados tem de estar sempre acessível e organizado, para a empresa poder consultá-los de forma simples, rápida e interativa. O sucesso de uma empresa está na análise da informação em seu poder. Anteriormente, para analisar os resultados laboratoriais utilizavam várias ODBC's, o que tornava a informação insegura e trabalhada de forma heterogénea. Por forma, a dar resposta a esta necessidade foi desenvolvido um cubo *OLAP* (*Online Analytical Processing*), denominado QUALIS. Este BI (*Business Intelligence System*) é transversal às 8 UI's de rolhas, que compila a informação presente neste *software* e no *software SPC*, utilizado para monitorização dos processos produtivos.

1.2.2 Processo produtivo

No processo produtivo o ideal é que todos os produtos estejam dentro das especificações, ou seja, não existam não conformidades. O departamento da Qualidade tem um papel preponderante na conformidade do produto, tendo por base as especificações do produto e os requisitos de cliente. O valor do produto está inevitavelmente relacionado com a sua qualidade, pois através desta o produto alcança prestígio e satisfação do cliente, que influenciará o sucesso da empresa.

Dois exemplos de processos produtivos da Amorim & Irmãos, S.A. são:



Figura 1.3: Processo Produtivo de Rolha Natural

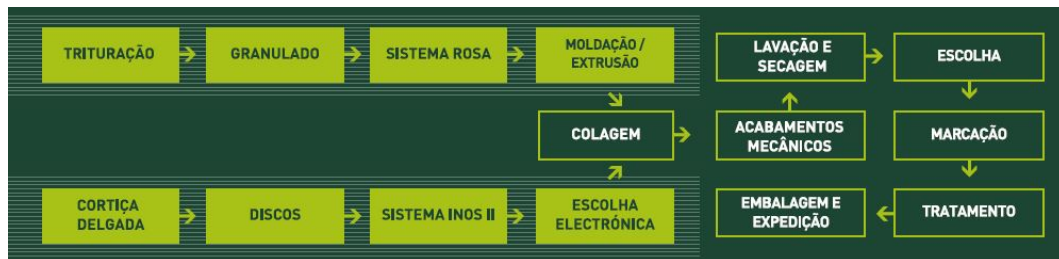


Figura 1.4: Processo Produtivo de Rolha Aglomerada com discos

Por forma, a responder aos requisitos dos clientes a melhoria da qualidade tem de ser um processo sistemático, para que o produto/processo esteja em constante actualização e/ou adaptação. As 7 ferramentas fundamentais para o controlo de qualidade são: Fluxograma, Diagrama Causa- Efeito (Espinha-Peixe), Folha Verificação, Diagrama de Pareto, Histograma, Diagrama de Dispersão e Cartas de Controlo.

Na empresa, durante o processo produtivo, são retiradas amostras para o controlo do processo, o resultado destes são registados no *software SPC* que é direccionado para a monitorização do processo, permitindo a análise estatística dos dados através de cartas de controlo. As cartas de controlo utilizadas atualmente na empresa são as cartas de Shewhart mais básicas, cartas $\bar{X} - R$, $\bar{X} - S$ e p .

1.3 Organização do relatório

No primeiro capítulo são apresentados os objetivos para a realização do estágio curricular e a descrição da empresa, os seus produtos, bem como processos produtivos.

No segundo capítulo são explicados os principais sistemas de informação que são utilizados pela empresa no departamento da Qualidade. Assim, neste capítulo serão apresentados os dois sistemas de informação, incidindo-se principalmente sobre o processamento analítico. Também neste serão introduzidos conceitos e fundamentos sobre cartas de controlo, em particular sobre as cartas de Shewhart e EWMA.

No terceiro capítulo são descritas as alterações que foram efetuadas nos sistemas de informações e o caso de estudo efetuado com base nas cartas de controlo apresentadas.

No quarto capítulo serão apresentadas as conclusões e propostas para futuros projetos.

Capítulo 2

Estado da Arte

2.1 Sistemas de Informação

Um sistema de Informação (SI) consiste na interligação de elementos, tais como: dados, equipamentos e pessoas, que armazenam, trabalham e retornam informação ao utilizador, que apoiam as organizações nas suas funções e/ou processos.

Os SI's distinguem-se em dois grupos:

- SI estratégico: são os que têm como finalidade a tomada de decisão através da implementação do plano estratégico da empresa;
- SI competitivo: são os que contemplam a melhoria de custos da empresa, num plano competitivo de mercado.

A partir do SI obtêm-se os dados estruturados, no entanto a tomada de decisão não é decidida pelo sistema, é sempre concretizada pelo agente de decisão, isto é, não podemos ignorar a componente humana. Com a constante evolução tecnológica e a enorme concorrência, para a empresa progredir necessita de tomar as decisões fundamentadas e num curto espaço de tempo. Uma solução para estes estudos quando existe um grande volume de dados é o uso de BI.

Cada laboratório regista no *ControLab*, para as diferentes fases do processo produtivo, os dados individuais de cada amostra¹ de todos os testes realizados às rolhas, apesar dos testes efetuados serem iguais nas UI's a arquitetura das bases de dados destes são diferentes para corresponder ao tipo de exigências dos vários produtos. Anteriormente ao QUALIS, cada laboratório para analisar os resultados laboratoriais, utilizava uma grande gama de ODBC. Neste capítulo serão explicados os dois SI's, incidindo-se principalmente sobre o processamento analítico.

¹A dimensão da amostra é recolhida em função da quantidade de rolhas rececionadas e produzidas.

2.1.1 *Online Transaction Processing*

Online Transaction Processing (OLTP) é o termo usado para caracterizar um SI capaz de aceder e recolher a informação e atualizá-la a partir de uma base de dados, igualmente designado por SI rotineiro.

Muitas organizações recorrem a SI's baseados em OLTP para as suas recolhas de dados, uma vez que estes têm a possibilidade de recolher a informação de vários repositórios de dados, traduzindo-se em consultas simples e rápidas, sem necessitar códigos de programação muito complexos para esta extração. Através do *Microsoft Office Excel* e linguagem *SQL (Structured Query Language)* é possível executar estas consultas sem trazer custos adicionais.

As ODBC's são consideradas um SI OLTP, dado que este processamento transaccional é um sistema meramente operacional, não servindo de apoio à tomada de decisão, isto porque consiste numa ligação a Sistemas de Gestão de Bases de Dados (SGBD) onde são introduzidos e armazenados os dados. O OLTP permite a consulta dos dados em bruto, sendo necessário o conhecimento da arquitetura da base de dados para que a extração da informação seja correta, estes assentam em modelos de Relação-Entidade normalizada, informação unidimensional.

As principais propriedades destes sistema são (da Silva, 2011):

- dados detalhados em bruto;
- atualização ao momento, sem histórico;
- processos repetitivos;
- análise de transações isoladas, conhecidas à *priori*;
- pequenas quantidades de informação.

2.1.2 *Online Analytical Processing*

Segundo Thomsen (1997) OLAP é “An acronym that stands for Online Analytical Processing. It is meant to contrast with OLTP. The key aspects are that OLAP is analysis-based and decision-oriented”. Um cubo OLAP é uma análise multidimensional, capaz de organizar e trabalhar dados de várias fontes de dados, onde ocorre a agregação e tratamento dos dados, de forma a organizar os dados para apoio à tomada de decisão. Contudo não apresenta qual a melhor decisão, nem quais as alternativas para a tomada de decisão. No entanto, tem a propriedade de interatividade, proporcionando análises que o SI transaccional não possibilita.

Os sistemas de processamento analítico têm como objetivo transmitir conhecimento ao longo do tempo, para isso recorre-se à análise dos registos históricos. Apesar desta característica de histórico, o OLAP não consiste numa base de dados, assim um termo que está interligado ao processamento analítico é *Data Warehouse* (DW).

2.1.2.1 Data Warehouse

O conceito de *Data Warehouse* assumiu maiores proporções através dos Sistemas de Suporte de Decisão (DSS), quando os SI's mais básicos (OLTP) utilizados pelas empresas já não respondiam às necessidades destes. Estes sistemas necessitam de bases de dados eficientes e avançadas para o apoio à decisão.

O DW não pode ser considerado como uma base de dados, segundo Dr. Inmon, denominado pai desta tecnologia, DW é “Subject oriented, time variant, non-volatile, integrate”. Em suma, a estrutura do DW é orientada para as análises de assunto, todas as mudanças dos dados são registadas ao longo do tempo, e têm a particularidade que uma vez lidos pelo DW nunca serão eliminados deste e este recolhe a informação de diversas bases de dados que permite uma análise tendo em conta diversas perspetivas. Com o decorrer dos tempos, a característica não-volátil do DW tem vindo a sofrer alterações ou adaptações através de arquiteturas de DW que distinguem as fontes de dados do SI e a apresentação dos dados (Dill, 2002).

Uma arquitetura de DW é imprescindível num sistema de processamento analítico. Este armazena toda a informação num único repositório, existindo o histórico com dados sumarizados, que são atualizados num período de tempo. A visão no DW é estruturada por vários assuntos, para ir de encontro às várias perspetivas dos utilizadores, abordados por camadas multidimensionais. Dependendo das arquiteturas do DW este pode ser constituído por *Data Marts*.

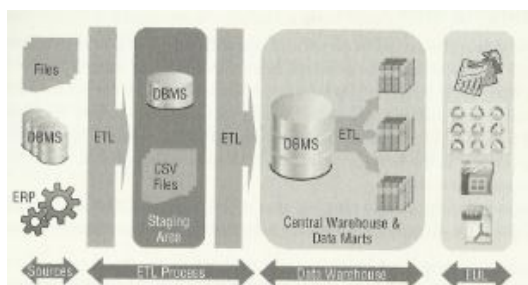


Figura 2.1: *Data Warehouse* com *Data Marts*

Fonte: (Bouman and Dongen, 2009)

2.1.2.2 Data Marts

Data Marts (DM) ou *Data Warehouse* Departamentais são os repositórios de menores dimensões que correspondem aos subconjuntos dos dados que separam os dados por assunto para responder às múltiplas análises pretendidas pelos utilizadores do sistema. Os dados podem estar repetidos em mais que um DM, isto é, pode existir interligações entre DM. Estas relações são estruturadas aquando do desenho da arquitetura da DW, as principais perspetivas de construção são (George, 2012):

- *Top-down*: define-se a estrutura global de toda a DW, a partir da DW gerar-se os DM (defendida por Dr. Inmon).
- *Bottom-up*: começa pela construção de DM fazendo experiências e/ou protótipos e de seguida constrói-se a DW global (defendida por Dr. Kimball).

Em suma, o DM é um banco de dados com ferramentas analíticas para o tratamento de assuntos específicos, pormenorizados e com menor dimensão que DW.

Outra etapa de grande importância na construção de uma DW é o processo ETL (*Extracting, Transforming and Loading*), traduzido por extração, transformação e transporte, ilustrado na Figura 2.1.

2.1.2.3 Extração, Transformação e Transporte

O processo de ETL dos dados para o DW situa-se entre as bases de dados recolhidos na empresa e os dados tratados na DW (da Silva, 2011). Nesta fase ocorre a extração de todos os dados em bruto, tal e qual como estão nas bases de dados de recolha de informação, sendo transferidos para a chamada área de estacionamento de dados (DAS). Esta área serve de repositório dos dados brutos onde estes serão transformados. A extração de dados para a DAS é bastante semelhante ao processamento transacional, uma vez que se trata de uma extração isenta de alteração dos dados.

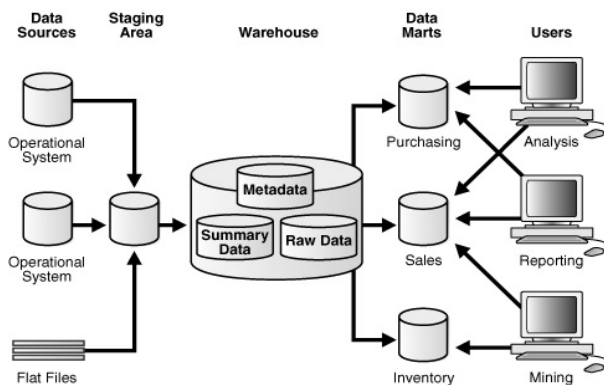


Figura 2.2: Ilustração de um *Data Warehouse* com área de estacionamento

Fonte: (Oracle, 2014)

A transformação de dados consiste na limpeza e preparação dos dados. Como visto anteriormente a DW poderá recolher a informação de diversas fontes de informação, onde nestas os dados podem estar estruturados de diferentes modos, assim é necessário proceder à eliminação das possíveis incoerências entre dados, para que seja assegurada a qualidade dos dados presentes na DW. Assim, é necessário proceder à filtragem de dados, selecionar apenas os campos que são essenciais para o

tratamento de dados. Tal como o levantamento dos possíveis relacionamentos entre campos, atribuição de nome aos campos e o tipo de informação deste (booleano, numérico, textual). Também aqui são efetuados os cálculos analíticos e agregações de dados.

Após todos os dados serem consolidados procede-se ao transporte destes dados da DAS para a DW, no entanto temos de ter em conta a dimensionalidade da DW.

2.1.2.4 Modelação Dimensional

A modelação dimensional é o que distingue fundamentalmente o processamento transacional e o processamento analítico, uma vez que é neste que definimos o nível de detalhe da DW. Os processamentos transacionais assentam em modelos Entidade-Relação que são unidimensionais. Não obstante, nos processamentos analíticos estes modelos simples não são os mais eficientes devido ao grande volume de dados e por esta razão os dados têm de ser estruturados e orientados para a consulta.

A estrutura consiste numa tabela principal e tabelas de dimensão inferiores.

A tabela principal é a tabela mais central, onde encontramos maior volume de informação centralizado, nestas estão as informações foco das análises do suporte à decisão, que correspondem à identificação das transações ou áreas de negócios. Em geral, estes recorrem a valores numéricos para a identificação destes. (Sarkis, 2001)

Nas tabelas de dimensões inferiores estão armazenadas menores quantidades de informação que contêm dados descritivos do negócio, utilizando representação textual. Estas ilustram hierarquias das áreas de negócio, que por sua vez, contêm as chaves primárias para que haja uma interligação entre a tabela principal e as tabelas de dimensão inferior.

As principais representações de modelos dimensionais são:

- Estrela: Tabela principal que está relacionada com um conjunto de tabelas dimensão.
- *Snowflake*: Tabela principal que está relacionada com tabelas dimensão, que são refinadas em tabelas de dimensão ainda mais reduzidas e normalizadas.

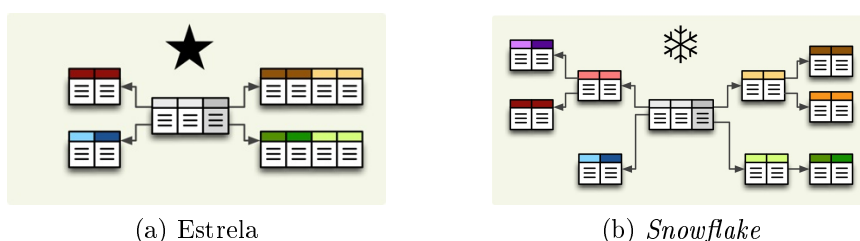


Figura 2.3: Ilustração de esquemas Estrela e *Snowflake*

Fonte: (Urbanek, 2012)

Neste contexto surge um conceito importante: granularidade. Este conceito corresponde ao nível de detalhe dos dados, quanto mais pequena é a granularidade do DW mais pormenorizados são os dados. Numa arquitetura em que a granularidade seja baixa, as consultas serão mais focalizadas numa área de negócio, o que exige uma maior capacidade de armazenamento de dados.

Existem algumas estratégias de armazenamento, que são: (Cavalcanti, 2012; Dill, 2002; Gorla, 2003; Sarkis, 2001)

- ROLAP - *Relational OLAP*: conjunto de interfaces de utilizador em que os dados e agregações são implementadas com bases de dados relacionais, recorrendo a linguagem SQL para a sua otimização nas consultas de dados;
- MOLAP - *Multidimensional OLAP*: os dados são armazenados em estruturas multidimensionais, assim as informações são agregadas em matrizes multidimensionais que facilitam o acesso rápido e direto;
- HOLAP - *Hybrid OLAP*: Resulta da combinação dos dois tipos de armazenamento anteriores, aliando a estrutura multidimensional e o ambiente relacional das bases de dados;
- DOLAP - *Dynamic OLAP*: armazenamento temporário dos dados, apenas para acesso rápido dos utilizadores;
- WOLAP - *Web OLAP*: usa-se plataformas independentes para acesso remoto à DW através de internet.

Igualmente relacionado com a dimensão da DW estão as operações dimensionais, que possibilitam uma flexibilidade e interatividade de recolha dos dados do cubo (Cavalcanti, 2012). As mais básicas são descritas na Tabela 2.1 e Figura 2.4:

Operação	Descrição
<i>Drill Down</i> ou <i>Roll Down</i>	Parte-se de um nível dimensional menos detalhado para o nível dimensional mais detalhado, podendo ser introduzida na análise mais dimensões, podendo atingir as variáveis das tabelas dimensão.
<i>Drill Up</i> ou <i>Roll up</i>	Agrega a informação para o nível hierárquico menos detalhado, resumando a informação (esta operação é o inverso do <i>drill down</i>).
<i>Slice</i>	Seleção de uma secção unidimensional do cubo.
<i>Dice</i>	Seleção de uma secção de duas ou mais dimensões do cubo; desta forma pode-se cruzar várias informações, ficando-se com o subcubo.
<i>Pivot</i> ou Rotação	Reorientar os eixos do cubo em análise, que pode corresponder a uma rotação em 2 ou 3 dimensões.

Tabela 2.1: Operações dimensionais em DW

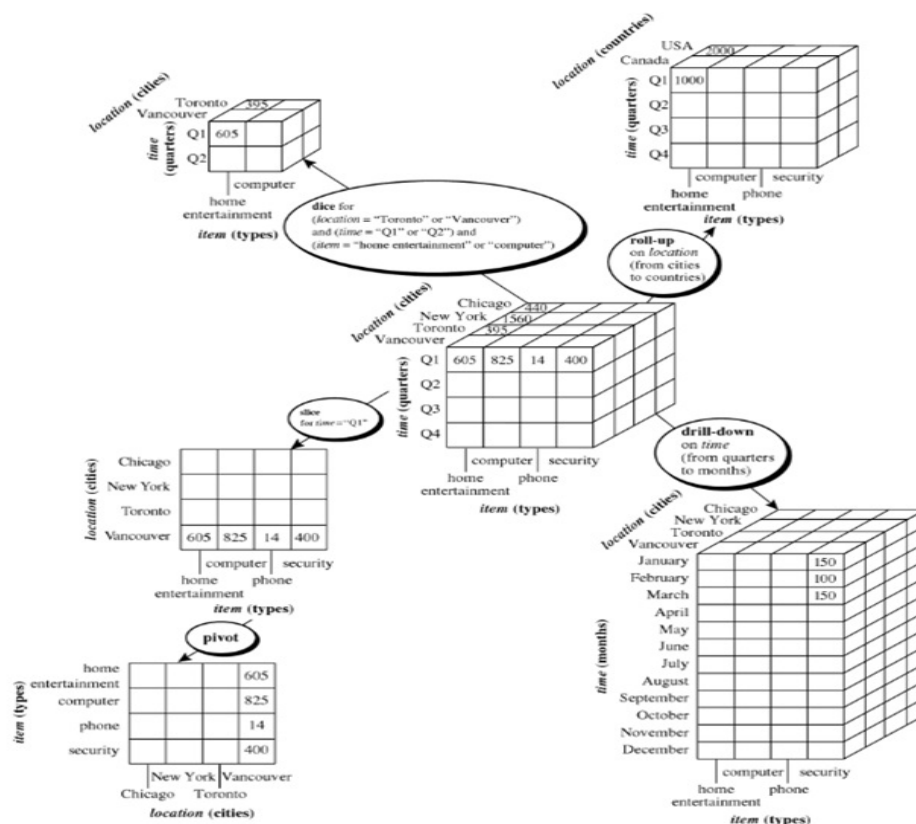


Figura 2.4: Ilustração de operações dimensionais
Fonte: (Cavalcanti, 2012)

2.1.2.5 Metadados

Os metadados são uma espécie de dicionário dos dados, isto porque são as informações de todos os dados, tais como: a origem destes, tipo, destino, as transformações a que são submetidos. Os metadados são cruciais na estrutura da DW, estes são assumidos como dados sobre dados.

Desde a fase de extração dos dados até à de transporte destes na DW é denominado *back-End* e a fase posterior a esta é denominado *front-End*. Em geral, na partição *front-End* os metadados são denominados por metadados técnicos, nestes estão definidas as origens de dados e seus relacionamentos e tipo de dados. Enquanto que na partição *back-End* são denominados metadados de negócios, nestes já estão especificadas as agregações e manipulações efetuadas aos dados para visualização dos utilizadores, tal como as operações dimensionais.

O *front-End* é a parte da DW onde se executam as traduções de todos os dados (ETL), enquanto na *back-End* é onde é mostrada toda a informação (interface com os utilizadores, ocultando toda a sofisticação do cubo OLAP). (Sarkis, 2001)

Como foi descrito na secção 2.1, o processamento OLAP é mais complexo e organizado que o processamento OLTP. O OLTP é considerado um processo do OLAP na fase de ETL.

2.2 Cartas de Controlo

Segundo Ribeiro and Caten (2012), a definição de Controlo Estatístico do Processo (CEP) é:

O controle estatístico do processo é um sistema de inspeção por amostragem, operando ao longo do processo, com o objetivo de verificar a presença de causas especiais, ou seja, causas que não são naturais ao processo e que podem prejudicar a qualidade do produto manufaturado. Uma vez identificadas as causas especiais, podemos atuar sobre elas, melhorando continuamente os processos de produção e, por conseguinte, a qualidade do produto final.

O CEP supõe um controlo do processo, mas é mais para além disso. O CEP permite-nos corrigir eventuais falhas no processo produtivo. Quando este já está a ser controlado é possível extrair informações que permitem otimizar o processo. Por tudo isto adquire-se um certo domínio sob processo, para além do controlo deste, obtendo portanto um processo produtivo robusto. O CEP não é usado apenas para a deteção de problemas, este também é fundamental para o aperfeiçoamento contínuo do processo. Em qualquer processo de produção o que se pretende é que os produtos sejam todos iguais e que estejam dentro das especificações, no entanto isto não se verifica existindo sempre uma certa variabilidade associada ao processo.

Estas variabilidades² podem ser classificadas segundo 2 tipos:

Causas	Descrição	Estado do Processo	Ação Corretiva
Aleatórias ou Comuns	Causas inerentes ao processo, tornando-se difíceis de eliminar.	Sob controlo estatístico.	O processo produtivo deverá decorrer normalmente.
Determinísticas ou Especiais	Causas consideradas falhas de operação.	Fora de controlo estatístico.	O processo produtivo deve ser interrompido, para a identificação e eliminação destas causas.

Tabela 2.2: Causas de variabilidade no CEP

²A variabilidade e suas causas foram estudadas por Dr. William E. Deming.

Uma ferramenta de CEP que permite a distinção entre as causas aleatórias e determinísticas é as cartas de controlo³, porque através desta representação gráfica facilmente se identificam as ocorrências de causas determinísticas. Estas representações pela sua fácil utilização e interpretação são um utensílio bastante utilizado na indústria. As cartas de controlo possibilitam que os operadores identifiquem qual o tipo de causa ocorrida num ponto de controlo (PC) para procederem à sua correção, em tempo real.

As cartas de controlo permitem a monitorização da qualidade de um processo produtivo. Estas são agrupadas com base nos dados disponíveis com que se pretende monitorizar a qualidade, em cartas para **atributos** ou para **variáveis**, conforme explicado em Esteves (2009). As cartas de controlo para atributos são usadas quando a finalidade não é “medir” as características da qualidade, mas sim, quando apenas são tidas em conta duas categorias, tais como: não conforme/defeituoso⁴ e conforme/não defeituoso.

Os tipos de cartas de controlo⁵ mais utilizadas para atributos são:

- **carta-*p***: análise da proporção/percentagem de produtos defeituosos;
- **carta-*np***: número de produtos defeituosos;
- **carta-*c***: número de defeitos, amostra de dimensão constante;
- **carta-*u***: número de defeitos por produto, amostra de dimensão variável.

As cartas de controlo para variáveis são utilizadas para a avaliação de dados mensuráveis que possam ser expressos como uma medida contínua. Estas cartas de controlo devem ser representadas por uma medida de tendência central e uma medida de dispersão (Montgomery, 1996). Estas contêm mais informação que as cartas anteriormente descritas, sendo usadas frequentemente as cartas para médias, desvios-padrão e amplitudes amostrais.

Para a elaboração destas são recolhidas sucessivas amostras⁶ para a inspeção dos produtos, sendo representadas graficamente através de uma medida estatística amostral adequada. Os resultados são confrontados com os limites de controlo, representados através:

³As primeiras cartas de controlo surgiram, em Maio de 1924, por Dr. Walter A. Shewhart (chamado pai do controlo estatístico moderno).

⁴Um produto defeituoso pode ter um ou vários defeitos, enquanto que um produto com um defeito não tem necessariamente de ser defeituoso. Um produto é considerado defeituoso quando não é aceitável.

⁵As cartas *p* e *np* baseiam-se na distribuição Binomial $Bi(n, p)$. Por sua vez as cartas *c* e *u* baseiam-se na distribuição Poisson.

⁶As amostras são classificadas segundo o período de recolha destas, se são recolhidas em intervalos de tempo fixo são denominadas por FSI (*Fixed Sampling Interval*), se forem em intervalos de tempo variáveis são denominadas por VSI (*Variable Sampling Interval*), consultar Gomes et al. (2010).

- Linha central (LC);
- Limites de controlo: linhas de controlo superior (LCS) e inferior (LCI);
- Limites de vigilância⁷: linhas de vigilância superior (LVS) e inferior (LVI).

Na Figura 2.5 está representado um exemplo de uma carta de controlo. Neste caso, a carta de controlo representada é uma carta para valores individuais, uma vez que a dimensão das amostras recolhidas é 1. Assim, no eixo vertical está indicada a característica de qualidade e no eixo horizontal está a indicação do número da amostra recolhida.

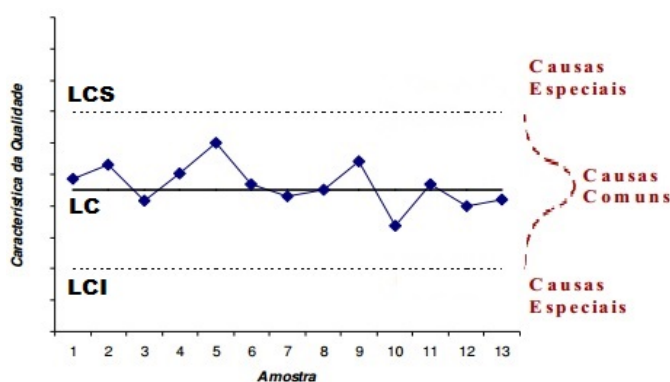


Figura 2.5: Exemplo de carta de controlo

Em suma, numa carta de controlo, no eixo das abcissas é representado o instante da recolha da amostra ou o número da amostra e no eixo das ordenadas é representada uma estatística amostral adequada para o parâmetro do processo que pretendemos monitorizar, por exemplo, média, desvio-padrão, amplitude.

Se os valores da estatística de controlo, ao longo do tempo, estiverem dentro dos limites de controlo o processo produtivo está sob controlo. A variabilidade destas amostragens dentro destes limites estão associadas a causas comuns ao processo. O processo produtivo é considerado instável, quando ocorrem causas especiais (estando fora de controlo) o que impossibilita prever o comportamento do processo. Estando perante um processo instável, é necessário interrompê-lo e proceder à identificação das causas determinísticas que ocorreram.

⁷Estes são optativos, mais usuais nas amostragens VSI.

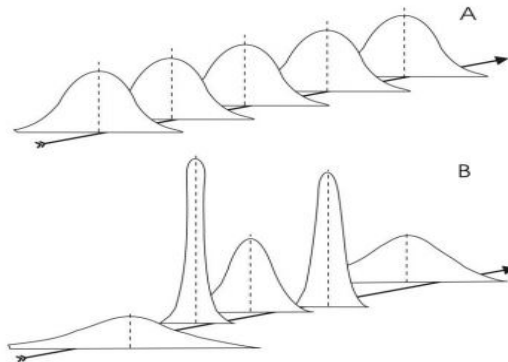


Figura 2.6: Processo estável versus Processo instável
Fonte: (Bonduelle, 2014)

Na Figura 2.6 (A) é visível um processo produtivo estável. Ao longo do tempo o processo produtivo tem uma média e desvio-padrão semelhantes, não existindo grandes variações, sendo previsível o comportamento no instante seguinte. Um processo produtivo estável pode estar fora de controle. Apenas é um processo produtivo estável sob controle estatístico quando os parâmetros do processo estão dentro dos limites de controle.

Na ilustração 2.6 (B) está-se perante um processo produtivo instável. Não é possível prever qual será o comportamento do processo produtivo no instante seguinte. Em cada instante, ocorrem “saltos” que poderão conduzir a amostragens fora dos limites de controle, estando perante causas especiais.

Nas cartas de controle em que todos os pontos se encontram dentro dos limites de controle, não é exatamente sinónimo que o processo esteja controlado. Desta forma, caso se verifique uma forma não aleatória da distribuição dos pontos na carta de controle é preciso analisar em maior detalhe a carta de controle, de forma a avaliar se o processo produtivo está efetivamente sob controle (Gomes et al., 2010). Assim, as cartas de controle podem ser interpretadas como um teste de hipóteses ao longo do tempo, ou seja:

$$\begin{cases} H_0 : \text{processo está sob controlo estatístico} \\ H_1 : \text{processo está fora de controlo estatístico} \end{cases} .$$

Para a realização do **teste de hipóteses** retira-se uma amostra aleatória da população que seja representativa desta, sobre a qual é calculada a estatística de teste que nos indica se devemos rejeitar ou não a hipótese nula (H_0).

Se H_0 é rejeitada quando é verdade, ocorre o erro tipo I. A probabilidade de ocorrência do erro tipo I é dado por

$$\alpha = P\{\text{ocorrer erro tipo I}\} = P\{\text{rejeitar } H_0 | H_0 \text{ ser verdade}\}, \quad (2.1)$$

também denominado **nível de significância**.

Se H_0 não é rejeitada quando é falsa comete-se o erro tipo II. A probabilidade deste erro ocorrer é dado por

$$\beta = P\{\text{ocorrer erro tipo II}\} = P\{\text{não rejeitar } H_0 | H_1 \text{ é verdade}\}. \quad (2.2)$$

Não é possível otimizar as duas probabilidades de erro em simultâneo. Assim, o objetivo é que o α seja baixo e consequentemente β seja alto.

2.2.1 Cartas de Controlo de Shewhart

As cartas de controlo de Shewhart foram as primeiras cartas a ser apresentadas, sendo as mais simples. Dr. Shewhart desenvolveu cartas de controlo para atributos e variáveis.

Este trabalho é focalizado nas cartas de controlo para variáveis. Estas cartas de controlo devem ser avaliadas com duas estatísticas descritivas: uma que traduza a localização e outra a dispersão dos dados. Assim, a carta de Shewhart para variáveis é um conjunto de duas cartas.

Neste trabalho serão analisadas as cartas: $\bar{X} - R$ (média e amplitude) e $\bar{X} - S$ (média e desvio-padrão). Segundo Montgomery (1996), as cartas de controlo $\bar{X} - S$ são usadas quando a dimensão da amostra (n) é variável ou quando a dimensão desta é elevada (superior a 10 ou 12). Dado que a sensibilidade da amplitude como estimador do desvio-padrão do processo produtivo diminui o seu desempenho com amostras elevadas. Uma vez que, o valor da amplitude apenas informa se os valores extremos da amostra estão próximos ou distantes, não sabendo o comportamento dos restantes dados amostrais. Caso uma amostra de elevada dimensão contenha um *outlier* a amplitude assumirá um valor elevado, enquanto que o desvio-padrão não será tão afetado. O desvio-padrão tem em conta todos os pontos da amostra, dando uma informação mais realista do que acontece ao longo do processo produtivo.

As cartas de controlo de Shewhart mais usuais são construídas tendo como pressupostos que as amostras são independentes e que a variável em estudo segue uma distribuição Normal (ou aproximadamente) $N(\mu_0, \sigma_0)$, μ_0 e σ_0 denotam o valor médio e desvio-padrão do processo sob controlo, respetivamente. Estes parâmetros podem ser conhecidos, ou não.

2.2.1.1 Coeficientes de assimetria e achatamento

A distribuição Normal é simétrica em relação à média. O achatamento da curva da distribuição dos dados está relacionado com o desvio-padrão da distribuição. Quanto menor for o desvio-padrão dos dados, menor é a dispersão da distribuição. Desta forma, é possível retirar-se conclusões sobre a possível normalidade dos dados, através do cálculo do **coeficiente de assimetria** (*skewness*) e do **coeficiente de achatamento/curtose** (*kurtosis*) dos dados.

Caso, os dados sejam simétricos tem-se um coeficiente de assimetria igual a zero. Dados que verificam uma distribuição com um coeficiente de assimetria negativa indica que os dados estão concentrados na “cauda” direita da distribuição, enquanto que o coeficiente positivo indica um enviesamento dos dados em relação à média para a esquerda.

A distribuição Normal tem um coeficiente de achatamento igual a três, segundo Pearson⁸. Dados que verifiquem um coeficiente de achatamento inferior verificam uma forma da distribuição com um “pico” mais baixo que a distribuição Normal.







Forma	Coefficiente Simetria	Forma	Coefficiente Achatamento
	Negativo		Inferior a Três
	Nulo		Três
	Positivo		Superior a Três

Figura 2.7: Ilustração de coeficientes de assimetria e achatamento

Em suma, dados que seguem uma distribuição Normal perfeita têm um coeficiente de assimetria nulo e coeficiente de achatamento igual a três. Dados que estejam desviado destes valores referência indicam que os dados não seguem uma distribuição Normal. No entanto, estas conclusões devem ser fundamentadas com recurso a testes de Normalidade.

2.2.1.2 Testes de Normalidade

Estes são testes de hipóteses usados para determinar se uma distribuição é proveniente de uma distribuição Normal. Os testes mais usuais são: Kolmogorov-Smirnov e Shapiro-Wilk.

O teste **Kolmogorov-Smirnov** é usado quando se testa distribuições contínuas, que é traduzido pelo teste de hipóteses:

$$\begin{cases} H_0 : F(x) = F_0(x), \forall x \\ H_1 : \exists x, F(x) \neq F_0(x) \end{cases} .$$

onde $F(x)$ é a distribuição seguida pelos dados e $F_0(x)$ é a distribuição empírica associada aos dados. Assim, quando se avalia a Normalidade dos dados, $F_0(x)$ é a distribuição Normal.

⁸Alguns autores e *softwares* optam para o cálculo da curtose o coeficiente excesso de curtose, que corresponde à subtração de três unidades ao coeficiente Pearson. Assim, pelo excesso de coeficiente de achatamento a distribuição Normal assume o valor nulo.

Para o cálculo da estatística de teste ordena-se a amostra (de tamanho n): x_1, \dots, x_n , de seguida são calculadas $F_n(x_i)$ e $F(x_i)$, para todos os elementos da amostra. Se H_0 é verdade, a estatística de teste é obtida por:

$$D = \max(\max |F_n(x_i) - F(x_i)|, \max |F(x_i) - F_n(x_{i-1})|).$$

O **teste Shapiro-Wilk** é exclusivamente usado para avaliar a Normalidade da distribuição de uma amostragem, sendo mais específico que o teste de Kolmogorov-Smirnov, daí mais robusto (Razali and Wah, 2011).

Seja x_1, x_2, \dots, x_n a amostra a testar e $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ a amostra ordenada, a estatística de teste de Shapiro-Wilk é dada por:

$$W = \frac{b^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2}. \quad (2.3)$$

onde b é dado em função da dimensão da amostra:

$$b = \begin{cases} \sum_{i=1}^{\frac{n}{2}} a_{n-i+1} (x_{n-i+1} - x_{(i)}), & n \text{ par} \\ \sum_{i=1}^{\frac{n+1}{2}} a_{n-i+1} (x_{n-i+1} - x_{(i)}), & n \text{ ímpar} \end{cases} \quad (2.4)$$

por sua vez a_{n-i+1} são constantes tabeladas, determinadas pelos valores esperados e matriz de covariância de distribuição Normal, que podem ser consultadas em Shapiro and Wilk (1965).

Associado ao teste de hipóteses tem-se sempre um **nível de significância** e **Região de Rejeição** do teste. O nível de significância é a probabilidade de rejeitar a hipótese nula quando esta é verdade. Um nível de significância de 5% é sinónimo que em 100 vezes rejeitamos 5 vezes a hipótese nula, quando esta deveria ter sido aceite. De modo geral, o nível de significância usado é $\alpha = 1\%$ ou $\alpha = 5\%$. A partir deste é calculada a Região de Rejeição.

Com base nesta Região de Rejeição e na estatística de teste é tomada a decisão de rejeitar ou não a H_0 . Se o valor da estatística de teste estiver contido na região de rejeição a hipótese nula é rejeitada. Enquanto, que se o valor na estatística de teste não estiver contido na região de rejeição não é rejeitada a hipótese nula, daí nada se pode concluir para o nível de significância estabelecido.

Também associado ao teste de hipóteses, está o conceito de $p - value$. Este é o menor nível de significância com que se pode rejeitar a hipótese nula. Desta forma, se $p - value < \alpha$, rejeita-se H_0 aceitando-se H_1 ; caso contrário não se rejeita H_0 .

Seja W_t a estatística de controlo associada à carta de controlo escolhida. Desta forma, o valor esperado da estatística de controlo é representada por $E(W_t)$ e a

variância por $Var(W_t)$. Em geral, os limites das cartas de controlo são definidos por

$$\begin{cases} LCS = E(W_t) + k\sqrt{Var(W_t)} \\ LC = E(W_t) \\ LCI = E(W_t) - k\sqrt{Var(W_t)} \end{cases} . \quad (2.5)$$

Em geral, é usado o $k = 3$ sendo estas cartas designadas por cartas $3 - \sigma$. Nestas cartas dentro dos limites de controlo situam-se cerca de 99.73% da variação de causas comuns, devido ao pressuposto dos dados seguirem uma distribuição Normal (ou aproximadamente).

Os **pontos das cartas de controlo** são representados através da estatística amostral mais adequada para o parâmetro do processo em estudo. Seja (x_1, x_2, \dots, x_n) uma amostra de dimensão n recolhida para inspeção.

Quando se desenha uma carta de controlo de média, carta \bar{X} , tem-se que o ponto referente à amostra anterior é dado por

$$\bar{X} = \sum_{j=1}^n \frac{x_j}{n}. \quad (2.6)$$

No caso das cartas R , o ponto da carta de controlo relativo à amostra anterior retirada para inspeção, é representado pela amplitude da amostra:

$$R = \max(x_j) - \min(x_j), \quad j = 1, \dots, n. \quad (2.7)$$

Por sua vez, na carta de controlo S a amostra é representada por

$$s = \sqrt{\frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n - 1}}. \quad (2.8)$$

Os **limites de controlo** para as diferentes cartas de controlo serão calculados, tendo por base, se são ou não conhecidos os parâmetros da distribuição Normal, como explicado de seguida.

Se os **parâmetros do processo são conhecidos sob controlo**, são usados diretamente para calcular os limites das cartas de controlo.

No caso da carta de controlo para a média ($W_t = \bar{X}$), têm-se que $E(\bar{X}) = \mu_0$ e $Var(\bar{X}) = \frac{\sigma_0^2}{n}$, sendo os limites de controlo para a carta $3 - \sigma$ para a média, quando os parâmetros sob controlo são conhecidos, são dados por

$$\begin{cases} LCS = \mu_0 + 3\frac{\sigma_0}{\sqrt{n}} = \mu_0 + A\sigma_0 \\ LC = \mu_0 \\ LCI = \mu_0 - 3\frac{\sigma_0}{\sqrt{n}} = \mu_0 - A\sigma_0 \end{cases} . \quad (2.9)$$

onde A é uma constante tabelada que varia em função do tamanho da amostra (n).

No caso da carta de amplitudes, têm-se que o valor esperado da amplitude é dada por $d_2\sigma_0$, e o desvio-padrão da amplitude é dado por $d_3\sigma_0$, sendo d_2 e d_3 , valores constantes dependentes de n , tabelados no caso de dados normais, e σ_0 é o desvio-padrão do processo sob controlo.

Aqui a estatística de controlo é a amplitude, assim em 2.5 W_t passa a ser R . Pela análise anterior $E(R) = d_2\sigma_0$ e $Var(R) = d_3\sigma_0$, aplicando estas expressões em 2.5, obtêm-se os limites de controlo da carta R quando os parâmetros do processo sob controlo são conhecidos.

Agora para o cálculo das linhas para a carta S é necessário estimar o desvio-padrão de S . Quando os dados seguem uma distribuição aproximadamente Normal tem-se que um estimador de S é $c_4\sigma_0$, onde c_4 é uma constante que varia em função do tamanho da amostra n , tendo-se assim $E(S) = c_4\sigma_0$. O desvio-padrão de S é dado pela raiz quadrada da variância de S , como $Var(S) = E(S^2) - (E(S))^2 = \sigma_0^2 - c_4^2\sigma_0^2$, obtêm-se $\sigma_S = \sigma_0\sqrt{1 - c_4^2}$.

Colocando em 2.5 os dados anteriores, obtêm-se os limites de controlo da carta S quando os parâmetros do processo sob controlo são conhecidos.

Se os **parâmetros da distribuição Normal não forem conhecidos sob controlo**, é necessário estimá-los. Para isso, são recolhidos k subgrupos de dimensão n , retirados quando o processo está sob controlo, usados para estimar μ_0 e σ_0 . Segundo Oliveira et al. (2013), para a estimação deve-se ter 20 a 25 amostras. Sejam as k amostras dada pelos valores $(x_{11}, x_{12}, \dots, x_{1n}), \dots, (x_{k1}, x_{k2}, \dots, x_{kn})$.

O melhor valor para a estimação da média do processo, é dado pela média da média dos subgrupos, que é calculada por

$$\bar{\bar{X}} = \sum_{i=1}^k \frac{\bar{X}_i}{k}. \quad (2.10)$$

onde a média aritmética do subgrupo i é dado por

$$\bar{X}_i = \sum_{j=1}^n \frac{x_{ij}}{n}. \quad (2.11)$$

Desta forma, $\bar{\bar{X}}$ é um estimador do valor esperado da média do processo ($\hat{\mu}_0 = \bar{\bar{X}}$).

A média da amplitude do subgrupo i é

$$R_i = \max(x_{ij}) - \min(x_{ij}), \quad j = 1, \dots, n, \quad (2.12)$$

sendo a média das amplitudes dos k subgrupos usados para a estimação dos parâmetros obtida por

$$\bar{R} = \sum_{i=1}^k \frac{R_i}{k}. \quad (2.13)$$

Quando a estatística amostral adequada para a monitorização do parâmetro é a amplitude, têm-se que o estimador do valor esperado da amplitude é \bar{R} .

O desvio-padrão associado ao subgrupo i é dado por

$$s_i = \sqrt{\frac{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}{n-1}}, \quad (2.14)$$

sendo, a média dos desvios-padrão dos subgrupos obtido por

$$\bar{S} = \sum_{i=1}^k \frac{s_i}{k}, \quad (2.15)$$

o qual é um estimador enviesado para o desvio-padrão do processo.

Como visto anteriormente, o valor médio do processo sob controlo pode ser estimada por 2.10. Quanto ao desvio-padrão do processo sob controlo, pode ser estimado com recurso às medidas de dispersão calculada a partir dos k subgrupos recolhidos para o efeito, 2.12 e 2.14.

A estimativa de σ_0 tendo em conta a amplitude dos subgrupos é dada por

$$\hat{\sigma}_0 = \frac{\bar{R}}{d_2}. \quad (2.16)$$

onde d_2 é constante tabelada, calculada em função da dimensão das amostras (n).

A estimativa de σ_0 tendo em conta o desvio-padrão dos subgrupos é dada por

$$\hat{\sigma}_0 = \frac{\bar{S}}{c_4}. \quad (2.17)$$

onde c_4 é uma constante tabelada que varia em função do tamanho da amostra n .

A carta para a média quando os parâmetros do processo sob controlo são estimados, é calculada recorrendo a estimativas obtidas através dos estimadores apresentados em 2.10 e a 2.16 ou 2.17. As cartas de controlo devem ser analisadas com uma medida de localização e uma medida de dispersão. Assim, no cálculo da carta \bar{X} é usado 2.16 quando esta é analisada em conjunto com carta R (carta $\bar{X} - R$). Quando analisada a carta \bar{X} com a carta S é usado 2.17 (carta $\bar{X} - S$).

Substituindo os estimadores anteriores, em 2.5 obtêm-se os limites de controlo para as cartas \bar{X} para parâmetros do processo sob controlo desconhecidos.

Para o cálculo da carta de amplitudes, necessita-se do valor esperado de R (dado por 2.13) e o desvio-padrão da amplitude. Uma estimativa do desvio padrão da amplitude é $\hat{\sigma}_R = d_3\sigma_0$, onde d_3 é uma constante tabelada dependente de n . Para além disso, $\hat{\sigma}_0 = \frac{\bar{R}}{d_2}$, de onde se obtêm $\hat{\sigma}_R = d_3 \frac{\bar{R}}{d_2}$.

Os limites da carta R para μ_0 e σ_0 desconhecidos são obtidos através de 2.5 com a substituição das expressões anteriores.

No cálculo da carta S , têm-se que o valor esperado de S é dado por 2.15 e um estimador para $\hat{\sigma}_S$ é igual a $\sigma_0\sqrt{1-c_4^2}$. Como o desvio-padrão do processo pode ser estimado por 2.17, têm-se que $\hat{\sigma}_S = \frac{\bar{S}\sqrt{1-c_4^2}}{c_4}$.

Colocando os valores estimados, \bar{S} e $\hat{\sigma}_S$, em 2.5 fica-se com os limites de controlo da carta S , quando os valores de μ_0 e σ_0 são desconhecidos.

Como as cartas de controlo de Shewhart são as mais utilizadas na indústria, apresentam-se em seguida um resumo dos diferentes limites destas cartas, o que facilitam a sua implementação.

Cartas de controlo $\bar{X} - R$

Parâmetros do processo fornecidos
Carta de Controlo \bar{X}

$$\begin{cases} LCS = \mu_0 + A\sigma_0 \\ LC = \mu_0 \\ LCI = \mu_0 - A\sigma_0 \end{cases} \quad .(2.18)$$

Carta de Controlo R

$$\begin{cases} LCS = D_2\sigma_0 \\ LC = d_2\sigma_0 \\ LCI = D_1\sigma_0 \end{cases} \quad .(2.19)$$

Parâmetros do processo estimados
Carta de Controlo \bar{X}

$$\begin{cases} LCS = \bar{\bar{X}} + A_2\bar{R} \\ LC = \bar{\bar{X}} \\ LCI = \bar{\bar{X}} - A_2\bar{R} \end{cases} \quad .(2.20)$$

Carta de Controlo R

$$\begin{cases} LCS = D_4\bar{R} \\ LC = \bar{R} \\ LCI = D_3\bar{R} \end{cases} \quad .(2.21)$$

Cartas de controlo $\bar{X} - S$

Parâmetros do processo fornecidos
Carta de Controlo \bar{X}

$$\begin{cases} LCS = \mu_0 + A\sigma_0 \\ LC = \mu_0 \\ LCI = \mu_0 - A\sigma_0 \end{cases} \quad .(2.22)$$

Carta de Controlo S

$$\begin{cases} LCS = B_6\sigma_0 \\ LC = c_4\sigma_0 \\ LCI = B_5\sigma_0 \end{cases} \quad .(2.23)$$

Parâmetros do processo estimados
Carta de Controlo \bar{X}

$$\begin{cases} LCS = \bar{\bar{X}} + A_3\bar{S} \\ LC = \bar{\bar{X}} \\ LCI = \bar{\bar{X}} - A_3\bar{S} \end{cases} \quad .(2.24)$$

Carta de Controlo S

$$\begin{cases} LCS = B_4\bar{S} \\ LC = \bar{S} \\ LCI = B_3\bar{S} \end{cases} \quad .(2.25)$$

$A, A_2, A_3, c_4, d_2, B_3, B_4, B_5, B_6, D_1, D_2, D_3, D_4$ são constantes tabeladas em função da dimensão dos subgrupos para modelo normal.

As cartas de controlo de Shewhart são de fácil construção e interpretação, no entanto é necessário ter cuidado nas conclusões retiradas. Em alguns casos, apesar das amostragens encontrarem-se dentro dos limites de controlo, o comportamento destas pode indiciar problemas no processo produtivo. Os sete critérios a ter em conta para que o processo produtivo esteja fora de controlo, foram elaborados por Dr. Shewhart que originaram a norma ISO 8258 - *Shewhart Control Charts*, como apresentado em Oliveira et al. (2013); Silveira (2013). Em Montgomery (1996) são apresentadas 10 regras para a interpretação desta temática nas Cartas de Controlo de Shewhart.

As cartas de controlo de Shewhart são boas para uma primeira fase do CEP. No entanto, quando já existir algum controlo sobre o processo estas não são as mais eficazes, já que estas cartas não são sensíveis a pequenas variações. Uma alternativa às cartas de controlo de Shewhart são as cartas de médias móveis (MA). As cartas MA têm em conta várias observações anteriores, por conseguinte irão detetar mais depressa pequenas alterações do processo. As alternativas mais utilizadas são as cartas CUSUM (*Cumulative Sum*) e as cartas EWMA (*Exponentially Weighted Moving Average*). Neste trabalho serão estudadas as cartas EWMA.

2.2.2 Cartas de Controlo EWMA

As cartas de controlo EWMA (*Exponentially Weighted Moving Average*) foram introduzidas em 1959 por Dr. Roberts, têm por base a média móvel exponencialmente pesada das amostras. Estas cartas têm a particularidade de em cada instante o subgrupo mais antigo ser ignorado, em que cada subgrupo tem determinado peso associado, favorecendo-se o mais recente no cálculo da estatística de EWMA. Os principais benefícios destas são a utilização de todos os pontos amostrais em cada instante, sensíveis aos pequenos desvios e mais adequadas a observações individuais, já que não são tão vulneráveis à hipótese da Normalidade dos dados. Estas cartas são preferíveis em relação às cartas de Shewhart quando o objetivo é o cumprimento das especificações impostas.

Quando a dimensão do subgrupo é unitária ($k = 1$), a **estatística EWMA** é representada por (Roberts, 2000):

$$Z_i = \lambda W_i + (1 - \lambda)Z_{i-1}, Z_0 = \theta_0, 0 < \lambda \leq 1, i = 1, 2, \dots \quad (2.26)$$

onde W_i é a **estatística de controlo**, ou seja a observação no instante i ; θ_0 é o **parâmetro que se pretende estimar** e λ é a **constante de amortecimento**, correspondente ao peso que é atribuído às observações anteriores e recente. A deter-

minação da constante de amortecimento depende da memória da estatística EWMA, como explicado em Hunter (1986).

Esta estatística pode ser utilizada em problemas de previsão e modelação de séries temporais, como referenciado em Gomes et al. (2010).

Montgomery (1996) afirma que a carta EWMA é insensível à hipótese de normalidade, visto que a média móvel tem em conta todas as observações antigas e recentes. Se as observações são independentes com variância σ_0^2 , têm-se que a variância de Z_i é dada por (Lucas and Saccucci, 1990):

$$\sigma_{Z_i}^2 = \sigma_0^2 \left(\frac{\lambda}{2-\lambda} \right) \left[1 - (1-\lambda)^{2i} \right]. \quad (2.27)$$

Assim, os limites de controlo exatos da carta de controlo EWMA para valores individuais são:

$$\begin{cases} LCS = \mu_0 + L\sigma_0\sqrt{\frac{\lambda}{2-\lambda} \left[1 - (1-\lambda)^{2i} \right]} \\ LC = \mu_0 \\ LCI = \mu_0 - L\sigma_0\sqrt{\frac{\lambda}{2-\lambda} \left[1 - (1-\lambda)^{2i} \right]} \end{cases}. \quad (2.28)$$

onde L é uma constante. A largura dos limites de controlo depende de L , que por sua vez varia em função de λ . A constante λ deve tomar valores baixos para se detetar pequenas variações do processo, estudos indicam que valores incluídos no intervalo $0.05 \leq \lambda \leq 0.25$ obtêm bons resultados, sendo em geral usados os valores 0.05, 0.10 e 0.20.

Como indicado anteriormente, nas cartas de controlo de Shewhart normalmente são usadas as cartas 3-*sigma*, nas cartas de controlo EWMA também pode ser usado este conceito ($L = 3$), mas assim deve-se ter um valor de λ superior aos de referência. Para os valores de λ referência, a constante L deve verificar $2.6 \leq L \leq 2.8$ (Montgomery, 1996). Em Crowder (1987) são apresentados resultados para valores tabelados de λ e L para estas cartas com base no *Average Run Length*⁹ (ARL).

Outro cálculo possível para os limites de controlo para a carta de controlo EWMA para valores individuais são:

$$\begin{cases} LCS = \mu_0 + L\sigma_0\sqrt{\frac{\lambda}{2-\lambda}} \\ LC = \mu_0 \\ LCI = \mu_0 - L\sigma_0\sqrt{\frac{\lambda}{2-\lambda}} \end{cases}. \quad (2.29)$$

⁹O ARL traduz o número de amostras (ou observações) recolhidas até a carta emitir sinal fora de controlo. Este é calculado em função da probabilidade de ocorrer o erro tipo I, $ARL = \frac{1}{\alpha}$ (Crowder, 1987; Montgomery, 1996).

pois à medida que o número de observações aumenta, a expressão 2.27 aproxima-se de $\sigma_{Z_i}^2 = \sigma_0^2 \left(\frac{\lambda}{2-\lambda} \right)$, devido a $\left[1 - (1-\lambda)^{2i} \right]$ aproximar-se de 1 quando i aumenta. Desta forma 2.29 são denominados limites de controlo assintóticos.

Ao utilizar os limites de controlo assintóticos nas primeiras amostras existe o risco de se concluir que o processo esteja sob controlo quando na realidade não está. Por essa razão para os primeiros valores da carta de controlo os limites de controlo devem ser calculados com recurso a 2.28.

A carta de controlo EWMA também é usada para monitorização dos valores médios do processo produtivo. Quando se está perante subgrupos com dimensão superior a 1 ($k > 1$), a observação i altera-se para a média do subgrupo i , ou seja, substitui-se W_i por \bar{X}_i e σ_0 por $\frac{\sigma_0}{\sqrt{k}}$, nas equações anteriores. Nestes casos a estatística EWMA, quando o valor médio do processo é conhecido, é traduzida por

$$Z_i = \lambda \bar{X}_i + (1 - \lambda) Z_{i-1}, Z_0 = \mu_0, 0 < \lambda \leq 1, i = 1, 2, \dots \quad (2.30)$$

μ_0 é conhecido por valor alvo para a média do processo produtivo.

Na carta de controlo EWMA existe uma espécie de combinação entre as cartas de controlo de Shewhart e a carta de controlo CUSUM, como referido em Crowder (1987); Gomes et al. (2010). Considerando-se $\lambda = 1$ a equação 2.26 torna-se uma carta de controlo de Shewhart e quando λ aproxima-se de zero a equação 2.26 assemelha-se à estatística de CUSUM¹⁰.

Tal como acontece nas cartas de controlo de Shewhart, as cartas de controlo EWMA devem ser analisadas tendo uma medida de tendência central e uma medida de dispersão. Anteriormente, foi analisada a carta de controlo EWMA para a média, para a análise da localização dos dados. Existem diversas propostas de cartas de controlo para análise da dispersão dos dados, que auxiliem as cartas de controlo EWMA, tais como: a carta de controlo EWMS (*Exponentially Weighted Mean Square Error*) apresentada em Montgomery (1996) e a carta $\ln(S^2)$ apresentada em Costa and Rahim (2006).

2.2.3 Capacidade de Processo

O estudo da capacidade de um processo informa se o processo é capaz de cumprir as especificações impostas, enquanto que as cartas de controlo indicam se o processo tem sofrido alterações ao longo do tempo. A capacidade de processo é avaliada por índices que verificam se um processo cumpre as especificações estabelecidas, sendo um aliado às cartas de controlo. Assim, para a avaliação da capacidade do processo é necessário que o processo esteja estável e que a variável em análise tenha uma distribuição aproximadamente Normal, se usarmos os índices de capacidade mais usuais (Bonduelle, 2014).

¹⁰A estatística de CUSUM é $S_i = S_{i-1} + (W_i - \theta_0)$, $S_0 = 0$.

2.2.3.1 Capacidade de Processo para distribuições Normais

A capacidade do processo é essencialmente traduzida por dois índices (os mais usuais): C_p e C_{pk} . Ambos têm em conta o desvio-padrão do processo σ_0 , mas apesar disso são distintos e utilizados em situações diferentes.

O índice C_p é mais simples que C_{pk} , apenas tem em conta a variabilidade do processo. Desta forma, não é sensível às causas especiais, não tem em conta a localização do processo.

Determinado com a **variabilidade permitida** (diferença entre o limite de especificação superior (LSE) e Inferior (LEI), respetivamente) e a **variabilidade inerente** ao processo¹¹ (esta é a variabilidade que corresponde ao Intervalo de Tolerância, $IT = LCS - LCI$):

$$C_p = \frac{LSE - LEI}{IT}. \quad (2.31)$$

Um processo produtivo é capaz quando este tem a capacidade de cumprir as especificações impostas ao processo, desta forma a variabilidade inerente ao processo tem de ser menor que a amplitude do intervalo definido pelas especificações impostas ($C_p > 1$), senão um porção de produtos estará fora dos limites de especificação.

Tendo em conta os valores de referências usuais retiram-se as seguintes conclusões acerca da capacidade do processo (Montgomery, 1996):

Processo incapaz	$C_p < 1$
Processo aceitável	$1 \leq C_p < 1.33$
Processo capaz	$C_p \geq 1.33$

Tabela 2.3: Valores referência para índice C_p

Outra medida usada, equivalente ao C_p é o seu recíproco, denominado **Razão da Capacidade** que é expresso na seguinte percentagem: $R_c = \frac{1}{C_p} \times 100\%$. Esta traduz a percentagem da especificação que está presente no processo, quanto maior for a razão da capacidade pior é o comportamento do processo. $C_p = 1.33 \Leftrightarrow R_c = 75\%$ (PortalAction, 2015).

É preciso ter cuidado com o C_p , pois pode ocorrer este ser elevado e estar a ser produzida uma grande probabilidade de não conformes (np). Este índice não tem em consideração o valor nominal da especificação (T), isto é posso ter um processo com uma variabilidade menor que a amplitude do limites de especificação (C_p elevado), mas completamente fora dos limites de especificação.

¹¹A variabilidade inerente do processo, associada às usuais cartas 3-sigma, é $6\sigma_0$, assim resulta:
 $C_p = \frac{LSE - LEI}{6\sigma_0}$.

O índice C_{pk} é usado quando não é verificada a localização do processo no valor nominal da especificação, assim sendo este índice tem em conta o valor da média do processo. Este indica se o processo produtivo está ou não descentrado, para além da sua variabilidade. Desta forma, considera-se o índice C_p a **capacidade potencial do processo** e o C_{pk} a **capacidade real do processo** (Bonduelle, 2014).

Para uma carta 3 – sigma, o índice C_{pk} é calculado através da seguinte equação:

$$C_{pk} = \min \left\{ \frac{LES - \mu_0}{3\sigma_0}, \frac{\mu_0 - LEI}{3\sigma_0} \right\}. \quad (2.32)$$

onde μ_0 e σ_0 são a média e o desvio-padrão do processo, respetivamente.

Este índice pode assumir valores negativos, o que indica que o processo produtivo está fora das especificações impostas.

Como C_p traduz a variabilidade do processo e C_{pk} traduz a localização do processo, devem ser interpretados em conjunto. Existem as possibilidades de obtenção de índices C_p e C_{pk} :

- Baixos C_p e C_{pk} : A variação do processo é maior que a variabilidade permitida com uma distribuição não localizada no valor nominal, o que se traduz numa maior probabilidade de np .
- Altos C_p e C_{pk} : A variação do processo é menor que a variabilidade permitida com o processo produtivo localizado no valor nominal, os valores estarão mais localizados no valor nominal correspondendo a uma menor probabilidade de np .
- Alto C_p e Baixo C_{pk} : A variação do processo é baixa em relação à amplitude dos limites de especificação, com um processo deslocado do valor nominal.

No conceito da capacidade do processo, foi introduzido por Dr. Genichi Taguchi, a **Função de Perda de Taguchi**. Dr. Taguchi afirma que a qualidade é a perda à atividade após o produto ser expedido. Desta forma, a qualidade não está apenas relacionada com o produto mas com as perdas que este trará à empresa.

Assim, afirma que o processo que cumpra os limites de especificação poderá ter ainda custos associados, ou seja, se a característica de controlo de processo não estiver localizada no valor nominal está-se a ter custos. Quanto maior for o afastamento do valor nominal da especificação, maior é a sua perda/custo.

A Função Perda de Taguchi igualmente conhecida por índice C_{pm} , é expressa por:

$$C_{pm} = \frac{LES - LEI}{6S_{Cpm}} = \frac{LES - LEI}{6\sqrt{\frac{\sum_{i=1}^n (x_i - T)^2}{n-1}}}. \quad (2.33)$$

onde T é o valor nominal da especificação e $S_{C_{pm}}$ representa o desvio-padrão ao valor alvo/valor nominal de especificação.

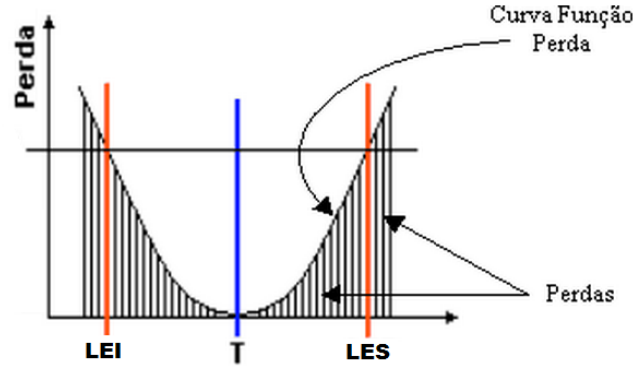


Figura 2.8: Função Perda de Taguchi
Adaptação: (Spectrum, 2010)

Há ainda um índice semelhante ao anterior mas que tem em consideração a localização do processo para o índice anterior:

$$C_{pmk} = \min \left(\frac{LES - \mu_0}{3\sqrt{\frac{\sum_{i=1}^n (x_i - T)^2}{n-1}}}, \frac{\mu_0 - LEI}{3\sqrt{\frac{\sum_{i=1}^n (x_i - T)^2}{n-1}}} \right). \quad (2.34)$$

Estes dois últimos índices não têm tanto em linha de conta as probabilidades de np , mas sim o afastamento do processo em relação ao valor nominal/alvo.

Devido à limitação dos índices de capacidade necessitarem que a variável siga uma distribuição normal são estudados novos índices de capacidade para quando ocorre esta limitação, tal como apresentado em Gonzalez and Werner (2009) e Khadse and Shinde (2009).

2.2.3.2 Capacidade de Processo para distribuições Não Normais

Como referido em Goswami and Dutta (2013) existem duas formas de lidar com os índices de capacidade para dados não normais:

1. Recorrer a transformações para adequar as distribuições Não Normais a distribuições Normais ou aproximadamente Normais e seguidamente usar os índices de capacidade usuais.
2. Usar índices de capacidade para distribuições Não Normais.

2.2.3.2.1 Transformações para Normalização de distribuições Não Normais

As transformações mais utilizadas atualmente para a normalização dos dados são: Box-Cox e Johnson¹², devido às suas fáceis implementações. Estas transformações são um pouco similares, mas a transformação de Box-Cox é utilizada para várias de valores positivos, enquanto que a transformação de Johnson é usada para a transformação de várias com valores negativos (aceitando também valores positivos).

Neste trabalho será utilizada a transformação de Box-Cox.

Transformação Box-Cox

A transformação de Box-Cox consiste na transformação dos dados originais X_1, X_2, \dots, X_n em Y_1, Y_2, \dots, Y_n segundo um parâmetro de transformação estimado, λ . A transformação dos dados é obtida por (Box and Cox, 1964):

$$\begin{cases} Y_i = \ln(X_i), & \text{se } \lambda = 0 \\ Y_i = \frac{X_i^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0 \end{cases} \quad (2.35)$$

O parâmetro de transformação λ pode tomar um vasto leque de valores, desta forma ficam cobertas diversas transformações matemáticas na transformação de dados.

As transformações de Box-Cox surgiram da família de curvas de Tukey (1957). As transformações mais antigas estão abrangidas neste método, como a raiz quadrada ($\lambda = \frac{1}{2}$) e raiz cúbica ($\lambda = \frac{1}{3}$), entre outras.

Osborne (2010) informa que as transformações de Tukey reduzem as assimetrias positivas, enquanto que para as assimetrias negativas estas transformações podem agravar as assimetrias. No entanto, a transformação Box-Cox não evidencia estes agravamentos.

Segundo Moraes et al. (2006), a transformação de Box-Cox é uma ferramenta eficiente na normalização de dados não Normais. Com esta transformação obtém-se observações normalmente distribuídas.

2.2.3.2.2 Índices de Capacidade para distribuições Não Normais

Na literatura, existem fundamentalmente três métodos que são empregues, que foram apresentados por: Clements (1989), Pearn and Chen (1997) e por Chen and Ding (2001).

¹²No *software* R existe o *package* 'Johnson' dedicado a esta temática (Fernandez, 2014).

Método Clements

Este método de cálculo dos índices de capacidade do processo, também é conhecido pelo método dos percentis. A base deste método assenta nas curvas de Karl Pearson e em três percentis da distribuição do processo.

Nos índices de capacidade usuais é tida em conta a variabilidade inerente ao processo como 6σ , que corresponde a uma probabilidade de erro de 0.27%. Desta forma, tendo em vista a distribuição Normal obtemos que os percentis usados para os cálculos deles são 0.135 e 99.865. Analogamente aos índices de capacidade usuais, Clements (1989) propôs o índice:

$$C'_p = \frac{LES - LEI}{F_{99.865} - F_{0.135}}. \quad (2.36)$$

onde F_p é o p -ésimo percentil¹³ da distribuição do processo.

Tal como nos índices de capacidade para dados normais, o índice C'_p , não tem em conta a localização do processo. Clements tal como anteriormente a partir dos índices de capacidade para dados normais, definiu o índice C'_{pk} que tem em conta o valor central da distribuição de dados. Por a distribuição dos dados não ser Normal, o valor central é calculado através da mediana (F_{50}), considerando a distância entre a mediana e os percentis (superior e inferior):

$$C'_{pk} = \min \left\{ \frac{LES - F_{50}}{F_{99.865} - F_{50}}, \frac{F_{50} - LEI}{F_{50} - F_{0.135}} \right\}. \quad (2.37)$$

Analogamente o índice C'_{pm} foi ajustado para:

$$C'_{pm} = \frac{LES - LEI}{6\sqrt{\left(\frac{F_{99.865} - F_{0.135}}{6}\right)^2 + (F_{50} - T)^2}}. \quad (2.38)$$

onde T é o valor nominal da especificação.

Para complementar o índice C'_{pm} tendo em atenção a localização do processo usa-se o índice C'_{pmk} :

$$C'_{pmk} = \min \left(\frac{LES - F_{50}}{3\sqrt{\left(\frac{F_{99.865} - F_{50}}{3}\right)^2 + (F_{50} - T)^2}}, \frac{F_{50} - LEI}{3\sqrt{\left(\frac{F_{50} - F_{0.135}}{3}\right)^2 + (F_{50} - T)^2}} \right). \quad (2.39)$$

Os índices de capacidade apresentados neste método se forem aplicados a distribuições Normais, são obtidos exatamente os mesmos valores que nos índices de capacidade para distribuições Normais. Ahmad et al. (2008) indica que os índices de

¹³O percentil F_p indica o valor para o qual existe $p\%$ valores inferior a este. É calculado tendo os valores ordenados crescentemente.

capacidade de Clements, não são os mais adequados para distribuições com notórias assimetrias. O estudo retratado em Gonçalves and Werner (2009) vai de encontro a esta conclusão.

Método Pearn and Chen

O método Pearn and Chen consiste numa adaptação do método de Clements. Este apenas difere no cálculos dos índices de capacidade que têm em conta a medida de localização: C'_{pk} e C'_{pmk} , dado que este método tem em consideração metade da distância entre percentis (superior e inferior). Assim,

$$C''_{pk} = \min \left\{ \frac{LES - F_{50}}{\frac{F_{99.865} - F_{0.135}}{2}}, \frac{F_{50} - LEI}{\frac{F_{99.865} - F_{0.135}}{2}} \right\}. \quad (2.40)$$

$$C''_{pmk} = \min \left\{ \frac{LES - F_{50}}{3\sqrt{\left(\frac{F_{99.865} - F_{0.135}}{6}\right)^2 + (F_{50} - T)^2}}, \frac{F_{50} - LEI}{3\sqrt{\left(\frac{F_{99.865} - F_{0.135}}{6}\right)^2 + (F_{50} - T)^2}} \right\}. \quad (2.41)$$

Atendendo a Gonçalves and Werner (2009), sabe-se que os índices pelo método de Pearn and Chen têm associada uma sensibilidade em identificar a distância entre o valor nominal da especificação e a mediana do processo: $C'_p < C'_{pk} < C'_{pm} < C''_{pm}$.

Método Chen and Ding

Este método é um pouco diferente dos restantes, uma vez que apenas considera um único indicador. Este índice é denotado por S_{pmk} , e tem como base de cálculo para a variabilidade do processo, a proporção de não conformes e a distância entre a média do processo e o valor nominal da especificação:

$$S_{pmk} = \frac{\phi^{-1} \left(\frac{1 + F(LSE) - F(LEI)}{2} \right)}{3\sqrt{1 + \left(\frac{\mu_0 - T}{\sigma_0} \right)^2}}, \quad (2.42)$$

F corresponde à função de distribuição acumulada do processo, ϕ é a função acumulada da distribuição Normal standardizada, μ_0 e σ_0 são média e variância do processo, respetivamente (Chen and Ding, 2001).

Comparativamente aos outros métodos anteriormente apresentados, seguindo a metodologia deste uma boa aproximação para o índice C_{pk} é dada pela seguinte expressão:

$$\frac{1}{3}\phi^{-1} \left(\frac{1 + F(LSE) - F(LEI)}{2} \right). \quad (2.43)$$

Ao utilizar o método de Chen and Ding o número de itens não conformes é maior que para os restantes índices de capacidade de processo.

Os índices de capacidade para distribuições Não Normais tem interpretações idênticas às dos índices de capacidade usuais. Em resumo:

- Índices de capacidade que apenas tem em conta a variabilidade do processo, quando assumem valores negativos é sinónimo que o processo produtivo não é capaz. Caso contrário as especificações impostas ao processo são alcançadas.
- Índices de capacidade que têm em conta a variabilidade e localização do processo caso sejam inferiores aos anteriores, a interpretação é que o processo não é centrado no valor nominal da especificação. Daí ser necessária a ação corretiva para o deslocamento da localização do processo. Se estes assumem valores negativos (o que é raro) é sinónimo que a medida de localização do processo produtivo está fora dos limites de especificação.

Capítulo 3

Trabalho desenvolvido no estágio curricular

Neste capítulo serão relatados e ilustrados os trabalhos efetuados ao longo do estágio curricular efetuado na Amorim & Irmãos, S.A.. Este teve como objetivo o desenvolvimento e tratamento dos dados da ferramenta BI da Qualidade da empresa, denominado QUALIS. O BI tem como fonte de dados dois *softwares* essenciais para o controlo da qualidade dos produtos: *ControLab* e *SPC*¹.

Neste projeto foram avaliados os campos dos SI's e propostas as melhorias necessárias, por forma a responder às necessidades dos utilizadores e para a existência de coerência entre informação dos SI's das 8 UI's da Amorim & Irmãos, S.A..

A nível de controlo estatístico de processo, foram estudadas cartas de controlo e índices de capacidade utilizados pelo *software* da empresa. Será apresentado um caso de estudo focado num PC do processo produtivo, denominado ROSA EVO presente na AI, recorrendo as ferramentas de controlo estatístico.

3.1 Sistemas de informação

No *software ControLab* são registados e armazenados os dados de todas as análises realizadas às rolhas, os valores recolhidos são confrontados com as especificações pré-definidas para a aprovação dos produtos.

Pelo que foi relatado em 1.2, cada UI tem as suas especificidades devido ao vasto portefólio de rolhas da Amorim & Irmãos, S.A.. Dada esta circunstância o *ControLab* apesar de ser semelhante em todas as UI's (este segue um esquema de bases de dados relacionais (Egitron, 2014)) a arquitetura das bases de dados são independentes, não existindo cruzamento/partilha de informação entre UI's por este sistema.

¹Os *softwares ControLab* e *SPC* foram desenvolvidos e são geridos pela empresa Egitron, enquanto que o QUALIS foi pela OSI - Sistemas Informáticos e Eletrotécnicos, Lda.

Não existe interligação deste SI com outro sistema para a alimentação dos seus dados mestres². Contudo possui ligação a equipamentos³ que efetuam a medição para as características controláveis definidas e transfere os dados para os boletins de controlo abertos no *software*. Simultaneamente, são gerados os boletins de controlo de qualidade dos produtos, que são direcionados aos clientes com os valores estatísticos dos diversos parâmetros de controlo (eventualmente valores individuais) obtidos nas amostragens de inspeção. Este *software* contempla uma *interface* com a plataforma *E-Supply*⁴, despoletando automaticamente o boletim de controlo para a plataforma, que posteriormente transmite ao cliente via correio eletrónico.

Os campos obrigatórios do boletim, à exceção da Quantidade, têm uma tabela com os valores pré-definidos que estes podem aceitar (estas tabelas são alimentadas pelos utilizadores). Como as UI's têm *softwares* independentes, para existir um SI que reúna a informação da qualidade das rolhas da Amorim & Irmãos, S.A. foi desenvolvido o BI - QUALIS. Os registos nas tabelas de cada *software* podem corresponder à mesma informação que existe noutra UI, mesmo assim caso esta informação esteja redigida de forma diferente, o cubo reconhecer-lhas-á como diferentes. Para poder exercer um paralelismo entre as informações das várias UI's é indispensável a uniformização dos campos, para que haja uma linguagem única que será exibida no *front-End* do BI.

3.1.1 Uniformização da informação

No início deste projeto já existiam algumas consolidações implementadas no cubo (tipo de produto analisado, categoria de relatório, tipologia de defeitos e designação de alguns fornecedores), que foram revistas, validadas ou adaptadas. Uma problemática associada a estas consolidações é a possibilidade de introdução de novas categorias nas diversas tabelas do *ControLab*. Estas novas introduções no BI serão consideradas como “Não Classificado”, o que implica a manutenção das consolidações.

Uma tabela fundamental na estrutura do *software* é a dos clientes/fornecedores. Ainda assim, esta tabela é problemática, pois pequenos erros de redação levam a criação de várias designações para a mesma entidade.

A informação das entidades chegam aos laboratórios através das notas de enco-

²Para a abertura de um boletim no *Controlab* é obrigatório o preenchimento de alguns campos, tais como: Categoria do boletim, Tipo de produto e respetiva Classe Visual, Calibre, Cliente/Fornecedor e Quantidade do lote analisado.

³Exemplo destes equipamentos são o MedCork (recolhe os dados relativos aos ensaios: dimensões, massa volúmica e humidade) e Aquaboy (aquisição da percentagem de humidade).

⁴É uma plataforma *online* onde ocorre um controlo das encomendas/expedições e reclamações/devoluções dos clientes. Neste cada encomenda tem associado o cliente/filial e a respetiva “entidade morada”.

menda, que são processadas nos sistemas: AS400⁵ ou SGPR⁶.

Nas notas de encomenda a identificação das entidades são apresentadas pela designação, marca do cliente ou identificação de embalagem. No entanto, para que uma entidade seja identificada de modo não ambíguo no AS400 é pela designação, código e dependência (que não são impressas nas notas de encomenda).

Na base de dados do AS400 para as 8 UI's existia um total de 32381 clientes e 10776 fornecedores. Nestes foram detetadas algumas incoerências, dado que pode existir “Cliente x” e “Cliente X”, que correspondem à mesma entidade, mas têm o código e/ou dependência diferentes. Foram detetadas 4808 designações iguais e 3315 códigos e dependência iguais. Isto é frequente acontecer quando a entidade se refere a uma filial, em que esta tem vários postos de entrega. A Amorim & Irmãos, S.A. fatura para a filial, mas o local de entrega corresponde às moradas dos clientes desta, daí ser necessário identificar várias direções com o mesmo cliente (filial). Para isto é criada uma nova entidade com a designação da filial e dependência diferente.

Assim, as analistas ao preencherem em *Controlab* a ficha da entidade, registam a designação impressa na nota de encomenda, que já pode trazer alguns erros. Para além disso, por vezes recorrem a abreviaturas nas designações.

A ficha da entidade do *Controlab* tem disponível vários campos, com estes o *software* emite os relatórios com o idioma apropriado, e através do grupo quando é aberto um boletim a associação dos ensaios necessários com o reconhecimento das especificações do grupo é automática.

Alguns laboratórios têm registado no código do *Controlab*, o código e dependência da entidade do AS400, que permite uma correspondência imediata entre as bases de dados.

Figura 3.1: Ficha entidade do *Controlab*

Inicialmente no cubo estavam consolidados unicamente os fornecedores de Rolha Natural das UI's AI, PK e VL, que totalizavam 247 entidades, que na realidade eram 197 entidades distintas. Esta consolidação não obedecia a critérios específicos para a sua nomeação. No total existiam 4485 entidades, distribuídas da seguinte forma:

⁵Esta ferramenta é utilizada nas 8 UI's. Nesta são geridas e criadas as encomendas internas/externas, referências para faturação, guias de transporte, gestão de *stock*, entre outras funcionalidades.

⁶O Sistema de Gestão de Produção de Rolhas é um SI mais “amigável” que o AS400, que contém uma grande *interface* de informação com o AS400.

UI	AD	AI	CK	DS	EQ	PK	TS	VL
Quantidade	1456	204	564	79	495	894	620	170

Tabela 3.1: Quantidade de entidades presentes no QUALIS, em função da UI

Para efetuar esta consolidação recorreu-se à base de dados do AS400. Para efetuar estas correspondências foi analisada a base de dados do AS400 e como cada UI efetuava os registos das entidades no *ControLab*.

Para a uniformização foi definido um plano de regras a seguir:

- Apenas letras maiúsculas;
- Remoção de siglas comerciais: SA, SL, SRL, SCRL, LDA, SAT, SCCL, SPA, SLU, SCL, EARL, EURL; Estas traziam imenso ruído nas correspondências: por vezes não eram consideradas ou eram escritas de várias formas. Por exemplo, Sociedade Anónima: “S.A.”, “SA.”, “SA”;
- Associadas às siglas comerciais surgiam “,”, “.”, “-”. Assim quando estes estavam no final do nome da entidade foram removidos;
- Por haver abreviatura de algumas palavras redigidas de diversas formas. Algumas transformações foram:

Referência	Adega	Bodega	Unipessoal	Cooperativa
Consolidação	ADEG.	BOD.	UNIP.	COOP.

- Eliminação de espaçamentos extra.

Assim, obtêm-se uma listagem de entidades final coesa e estas regras possibilitam que seja mantida sempre a mesma estrutura nas consolidações. Desta forma, as consolidações já existentes foram ignoradas, passando a respeitar estes critérios.

Foram aplicadas estas transformações à listagem das entidades presentes em AS400 e QUALIS. Após este tratamento de dados foram consolidadas 1689 entidades, tendo apenas em conta a designação presente no *ControLab*.

Na AI, PK e VL alguns dos fornecedores continham o código de AS400 na ficha da entidade do *ControLab* (Figura 3.1). Esta informação permitiu a identificação de mais 76 entidades.

Mesmo após este trabalho continuaram designações ambíguas, tendo sido necessário analisar caso a caso, nos registos históricos das UI's. Uma ferramenta que contribuiu nesta fase foi o *E-Supply*.⁷

⁷Não foi efetuado este trabalho logo de início, por esta ferramenta não contemplar o código e dependência do AS400.

No final desta consolidação as 4522 entidades presentes no QUALIS foram consolidadas em 3375 entidades. A taxa de redução foi aproximadamente 25.36%, a entidade que foi mais notória para esta redução corresponde a uma filial que possui 79 entidades diferentes no QUALIS (cerca de 1.75% das entidades do QUALIS).

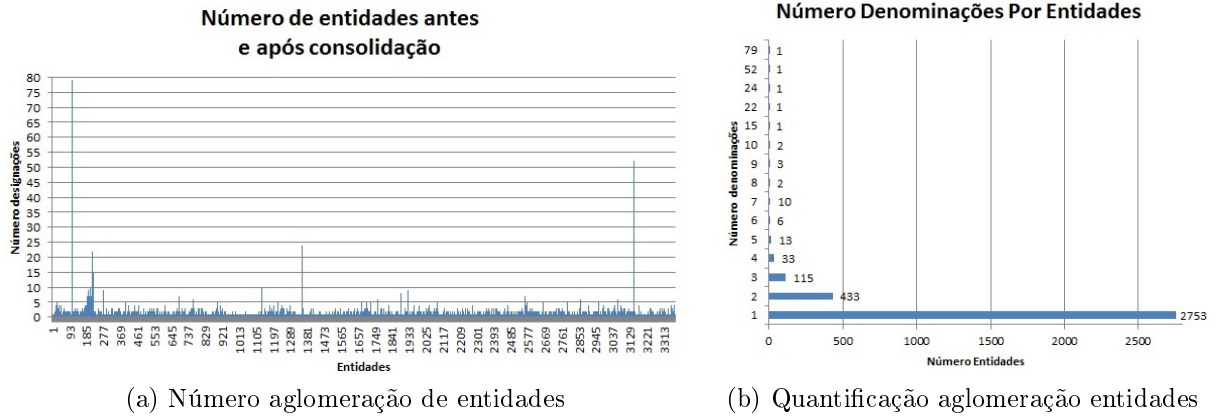


Figura 3.2: Impacto da Consolidação das entidades presentes no QUALIS

Este trabalho foi fundamental, pois num BI como o QUALIS toda a informação tem de estar correta para que não hajam análises incorretas. Num sistema OLTP o utilizador tem de conhecer bem a estrutura para conseguir extrair a informação pelos modelos de Relação-Entidade normalizados. Num sistema OLAP os utilizadores já encontram um ambiente mais interativo e intuitivo, que requer um grande trabalho de *backstage* para que a informação não seja fraudulenta.

3.1.2 Acompanhamento a fornecedores

A empresa para conquistar uma posição competitiva face ao mercado tem de conhecer profundamente as suas matérias-primas e produtos. Neste âmbito foi implementado na empresa um acompanhamento mais focalizado aos fornecedores de Rolha Natural, devido à problemática atual do TCA na rolha de cortiça. O TCA é um composto presente na Natureza responsável pelo gosto a rolha no vinho, no entanto este não traz nenhum risco para a saúde. Este composto não está somente associado à cortiça, pode estar presente na água, vinho, legumes. Já existe alguns mecanismos para prevenção da formação de TCA, não existindo ainda uma solução que garanta a completa ausência deste composto (APCOR, 2011).

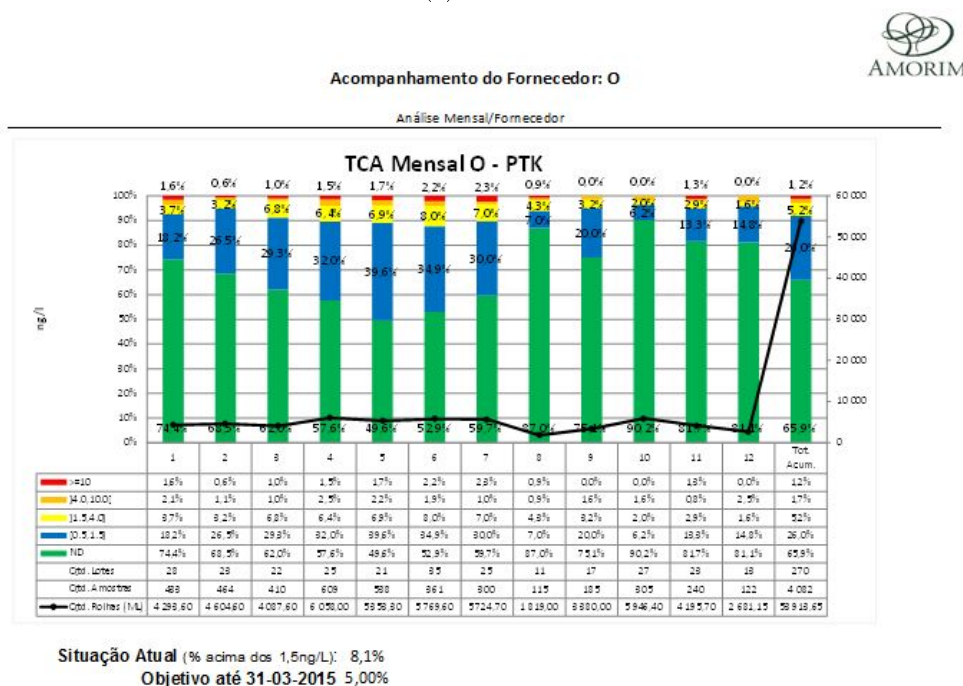
Foram selecionados 15 fornecedores das UI's AI, PK e VL, que são visitados mensalmente onde apreciam os seus fornecimentos, tendo em conta a quantificação de TCA dos fornecimentos. O composto é quantificado em ng/L , quando é detetada

quantidade abaixo de 0.5 ng/L é considerado ND (Não detetável), sendo considerado sensorialmente crítico desde 1.5 ng/L . Neste acompanhamento é tido em conta duas gamas intervalares para o enquadramento do TCA. Para o efeito foram criados ficheiros Excel com a extração dos dados via QUALIS, que contém programação com linguagem VBA (*Visual Basic for Applications*) para ser possível a compilação das várias consultas sobre o cubo no tratamento dos dados.

Numa primeira fase foi criado um *Report*⁸ onde é possível analisar os níveis de TCA obtidos no momento de receção da matéria-prima com a informação da quantidade adquirida (lotes, milhares de rolas (ML)) e número de amostras que foram efetuadas para a quantificação do composto.

Fornecedor	Mês	Qtd. Lotes	Qtd. Amostras	Qtd. Rolas	ND	0.5-1.5	1.5-4.0	4.0-10.0	10-16	% acima 1.5	Atualizar
O	1	28	433	4 293,60	74,4%	18,2%	3,7%	2,1%	1,6%	7,4%	
O	2	23	464	4 604,60	68,5%	26,5%	3,2%	1,1%	0,6%	5,0%	
O	3	22	410	4 087,60	62,0%	29,3%	6,8%	1,0%	1,0%	8,8%	
O	4	25	609	6 058,00	57,6%	32,0%	6,4%	2,5%	1,5%	10,3%	
O	5	21	538	5 353,30	49,6%	39,6%	6,9%	2,2%	1,7%	10,8%	
O	6	35	361	5 769,60	52,9%	34,9%	8,0%	1,9%	2,2%	12,2%	
O	7	25	300	5 724,70	59,7%	30,0%	7,0%	1,0%	2,3%	10,3%	
O	8	11	115	1 819,00	87,0%	7,0%	4,3%	0,9%	0,9%	6,3%	
O	9	17	185	3 380,00	75,1%	20,0%	3,2%	1,6%	0,0%	4,9%	
O	10	27	305	5 946,40	90,2%	6,2%	2,0%	1,6%	0,0%	3,6%	
O	11	23	240	4 195,70	81,7%	13,3%	2,9%	0,8%	1,3%	5,0%	
O	12	13	122	2 681,15	81,1%	14,8%	1,6%	2,5%	0,0%	4,1%	
O	Tot. Acum.	270	4 082	53 913,65	65,9%	26,0%	5,2%	1,7%	1,2%	8,1%	

(a) Tabela



(b) Gráfico

Figura 3.3: Análise Mensal / Fornecedores

⁸Devido a confidencialidade nas ilustrações deste, os nomes e códigos dos fornecedores estão identificados com letras e números, respetivamente.

Em cada mês é avaliada a situação atual do fornecedor (percentagem de TCA superior a 1.5 ng/L), no início ficou definido um objetivo para cada fornecedor que deveriam atingir até ao dia 31-03-2015 para este patamar. A evolução mensal encontra-se apresentada em forma de tabela e graficamente, como ilustrado na Figura 3.3. Esta informação é cedida aos fornecedores nas reuniões mensais.

Para adquirir conhecimento acerca dos fornecedores é importante poder comparar os fornecimentos. Para além desta análise individualizada de fornecedores, foi criada uma consulta *drill-up* sobre a anterior, ficando uma análise anual que permite retirar conclusões da situação global dos 15 fornecedores⁹. Esta também contempla uma análise detalhada e alargada por gamas de TCA. Na Figura 3.4 está a informação que é entregue aos fornecedores, assim a identificação é efetuada através do código de AS400 para ser preservada a identificação das várias entidades que têm vindo a ser acompanhadas.

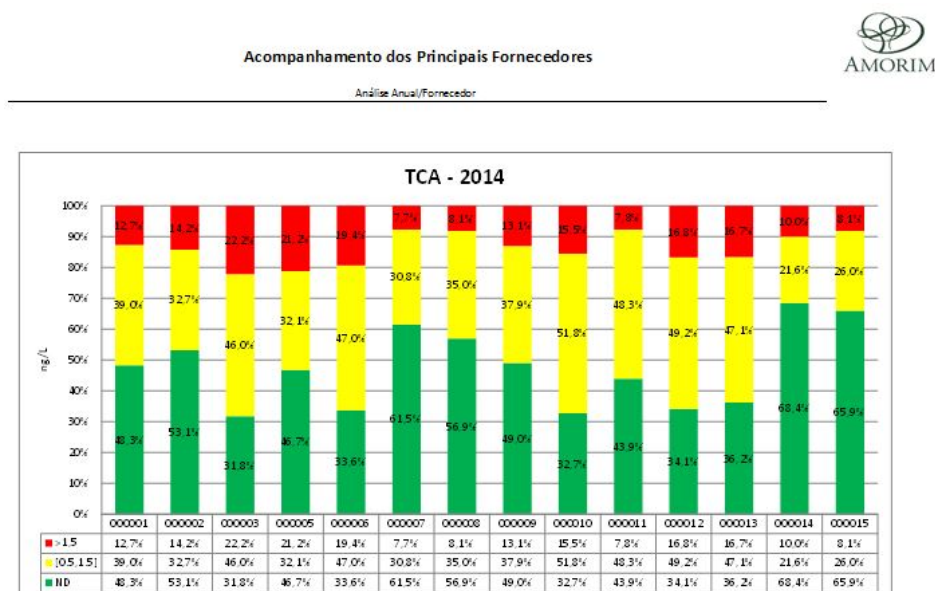


Figura 3.4: Análise Anual dos Principais Fornecedores

Nestas representações fica evidente quais são os piores fornecedores de Rolha Natural neste grupo, em termos de TCA. Nesta análise anual existe a hipótese de eliminar os fornecedores em que a sua situação atual se encontre superior a uma percentagem selecionada. Assim é possível avaliar quais os piores fornecedores, permitindo distinguir quais são os fornecedores que se deve manter e quais são prejudiciais para o negócio.

⁹Se um fornecedor abastecer UI distintas é possível especificar que UI está em análise.

Tal como é fulcral para a empresa conhecer os seus piores fornecedores, conhecer os seus melhores também o é. Numa segunda fase deste acompanhamento foi elaborado outro *Report* para extrair a informação de quais são os melhores fornecedores, chamados os “TOP 20”. Aqui são analisados os fornecimentos de Rolha Natural por entidades externas à Corticeira Amorim, S.G.P.S.,S.A..

Para além dos TOP 20 para cada UI, existe o TOP 20 Geral onde são apresentados os resultados de TCA dos 20 melhores fornecedores das três UI's. Um fornecedor é identificado também com a UI. Assim entidades que forneçam mais que uma UI, podem aparecer repetidamente associado às várias UI's que fornecem. Aqui a análise mensal efetuada aos 15 principais fornecedores pode ser efetuada para qualquer fornecedor, incluindo fornecedores internos. Desta forma, tem-se o enquadramento dos diversos fornecedores, incluindo a comparação da mesma entidade para fornecimentos em diferentes UI's.

Através destes *reports* a equipa de compra de rolhas Naturais tem bases para decidir como atuar com os seus fornecedores: quais necessitam de melhorar os seus resultados, quais os que deve manter e quais os que deve estar alerta ou cessar fornecimentos. A empresa com esta análise acrescenta valor aos seus produtos.

3.1.3 Quantificação de TCA

O processo de quantificação deste composto é complexo, que requer um equipamento de custos elevados denominado cromatógrafo. O departamento de Investigação e Desenvolvimento (I&D) da Amorim & Irmãos, S.A., pertencente à Corticeira Amorim, possui estes equipamentos onde são quantificadas as macerações das rolhas que são recolhidas nas diversas fases do processo produtivo, nas várias UI's. As metodologias para deteção de desvios sensoriais frequentemente usadas são *Dry-Soak* e *Soak*. O método *Dry-Soak* não é destrutivo para as rolhas, permitindo separar os desvios sensoriais pelos seus grupos. Todavia, não permite saber com exatidão quais as quantidades de composto presente nas rolhas. Por *Soak*, as rolhas têm de estar em maceração durante 24 horas em solução alcoólica em temperatura favorável para libertação do composto para ser possível a sua quantificação. Esta preparação segue para o I&D, onde é efetuada mais uma série de procedimentos para quantificação de TCA através dos cromatógrafos. Este ensaio é fundamental no controlo da qualidade do produto, estando associado a custos elevados e grandes volumes de tempo.

As rolhas que são transitadas entre UI's estão dentro dos padrões/especificações do grupo. Contudo devido a caderno de encargos que poderão ser diferentes das diretrizes do grupo, as rolhas são inspecionadas no momento da receção já tendo em conta o cliente para que serão destinadas. No momento de receção das rolhas se os lotes já vierem acompanhados com os dados obtidos na inspeção de expedição nas UI's fornecedoras poderão ser evitadas re-análises nomeadamente para a quantificação de TCA.

Nesta temática, existe a lacuna do não cruzamento de históricos entre as bases de dados do *ControLab* das várias UI's. Para resolução desta lacuna existe a possibilidade de extração da informação disponível na ferramenta de BI. O QUALIS tem o objetivo efetuar as análises estatísticas para a qualidade, mas também pode ser uma fonte de informação devido à possibilidade de extração dos valores individuais obtidos nos diversos ensaios.

Na AD ocorreu a tentativa de implementação de um ficheiro que exportasse a informação dos valores de TCA obtidos nas expedições da AI para a AD, para que à chegada dos lotes já tivessem esta informação evitando a recolha, preparação, e custos inerentes a este ensaio, por já ter sido realizado na AI. Essencialmente a AD recebe rolhas das UI's: AI, DS e EQ.

Aquando a criação deste ficheiro os registos na AI eram feitos de forma diferente da atual, com as alterações efetuadas nos registos foram perdidas as rastreabilidades estabelecidas. Neste estágio, foi levantada a rastreabilidade entre as ordens de fabrico das duas UI's. Para ser possível a consulta dos dados, este ficheiro foi adaptado. Os registos na AI para o TCA, atualmente são efetuados no *ControLab* e com recurso a ficheiro Excel. As amostras para análise de TCA quando são coletadas é-lhes associado uma referência (esta é registada em ambos os locais) e no ficheiro Excel é registado a ligação entre referência de amostra e a ordem de fabrico AI, sendo este ficheiro usado pela logística de saída como indicativo de lotes conformes ou não para o ensaio de TCA, que recolhe os valores individuais e estatísticas de TCA através do QUALIS.

Com recurso a uma consulta ODBC sobre o AS400 obtém-se os dados presentes nas guias de transporte dos lotes expedidos para a AD: produtos, quantidades, ordem de fabrico (OF) expedição AI, OF entrada AD). Através destas OF's é efetuada a pesquisa dos resultados de TCA, no cubo QUALIS (para lotes de *stock* da AI anteriores a esta alteração) e no ficheiro Excel da AI (este também foi adaptado para ser possível esta consulta).

Atualmente os dados de TCA recolhidos das amostragens efetuadas na AI estão disponíveis no ficheiro à entrada dos lotes na AD. Este trabalho também foi efetuado para os registos da DS¹⁰. Para a EQ esta recolha de valores de TCA não foi bem sucedida, dado que não registam todos os resultados obtidos no *ControLab* sendo registados em ficheiros Excel. Assim, para efetuar a rastreabilidade dos lotes envolve alguns ficheiros em rede, que torna inviável para as consultas em Excel.

Com estas adaptações ao ficheiro existente na AD, já não é necessário efetuar os ensaios de TCA nos lotes rececionados, havendo uma grande redução de ensaios e custos associados a este trabalho que ocorria em duplicado.

¹⁰O trabalho de rastreabilidade permitiu ter acesso aos valores dos restantes ensaios realizados para os lotes por estes estarem registados no *ControLab*.

3.1.4 Caraterística Volátil do QUALIS

Anteriormente ao QUALIS, em cada laboratório para analisar os resultados laboratoriais utilizavam ODBC's, que podia tornar-se num trabalho perigoso e muito penoso. O *Controlab* e *SPC* são de fácil utilização a nível dos utilizadores, mas as exigências e diversidade de opções que estes *softwares* dispõem, obrigam que a estrutura de bases relacionais seja bastante complexa. Desta forma, as ODBC's sobre estas bases de dados têm de ser cautelosas na sua criação, para que não haja confusão na identificação nas tabelas das chaves principais e secundárias.

O QUALIS permite fazer estas consultas, de uma forma muito mais cómoda, mais fiável e rápida. Este BI tem como finalidade a elaboração de análises estatísticas dos dados da Qualidade. Assim, temos de entender o que significa um registo ser considerado para tratamento estatístico.

O *SPC* é utilizado para a monitorização dos processos produtivos, sendo que todos os dados adquiridos por este *software* são contabilizados para o controlo do processo. Deste modo, todos são considerados para o tratamento estatístico, estando disponíveis no QUALIS.

No *Controlab* nem todos serão contabilizados para o tratamento estatístico. Este *software* tem essencialmente dois níveis hierárquicos: Relatórios e Ensaios. Os relatórios podem ter vários ensaios diferentes e/ou repetições dos ensaios. Cada um destes níveis tem associada uma *flag* denominada “Estatística”. Esta *flag* foi criada com o intuito de assinalar as informações que seriam utilizadas para as análises estatísticas, sendo então extraídas para o cubo OLAP. Pode haver casos em que os ensaios ou relatórios resultam de duplicação de registos já existentes, desta forma não podem ser contabilizados como carga de trabalho efetiva.

No caso de uma prestação de serviços que consista num retratamento de superfície, as dimensões das rolhas ficam inalteradas. No entanto as forças de extração destas serão alteradas. No final da prestação de serviços, o novo boletim de controlo terá de conter a informação atualizada das forças de extração, podendo o ensaio de dimensões ser uma duplicação do controlo efetuado na expedição, logo não pode ser contabilizado para tratamento estatístico por não ter sido realizado trabalho. Aqui fica ilustrada a importância da *flag* a nível do Ensaio.

A nível do relatório também é possível acontecer a não contabilização para tratamento estatístico. Por exemplo, um lote que repartindo-o satisfaz duas encomendas. Por tratar-se de um lote efetua-se um controlo, sendo o resultado deste válido para as duas encomendas (dois boletins). Então apenas um boletim será contabilizado para tratamento estatístico, e o outro que é repetição do anterior, terá a sua *flag* da estatística desativada. Um relatório que não conte para tratamento estatístico, nenhum dos seus ensaios contará para tratamento estatístico.

Devido às estruturas dos *softwares* utilizados pela Qualidade, o QUALIS segue um esquema normalizado *snowflake*, ver ilustração 2.3b. Devido à *flag* “Estatística”, nem toda informação armazenada nos *softwares* está visível no QUALIS.

O QUALIS é um BI baseado em tecnologia OLAP. Assim todos os dados são lidos das bases de dados e armazenados em histórico. Desta forma, os relatórios que não contam para estatística estão identificados no BI e têm a eles associados a indicação da não ativação da *flag*. O cubo tem o campo denominado “ShowInStat”, estando este desativado ao nível do relatório, tem-se a informação mais básica dos relatórios que não contam para tratamento estatístico. Esta informação é importante, pois permite ter a percepção da quantidade de rolhas que foram expedidas em confronto com a quantidade de rolhas que efetivamente foram analisadas (relatórios com *flag* ativada).

Nestas análises foi identificada uma discrepância de valores entre os relatórios presentes nas bases de dados do *Controlab* e o QUALIS, bem como a quantidade produzida. Aqui foi necessário efetuar um paralelismo a vários níveis hierárquicos para o reconhecimento das incoerências entre os SI's. As quantidades de relatórios registados no QUALIS eram sempre superiores, o que é contraproducente. Dado à dessincronização¹¹ da atualização do cubo, o que deveria acontecer era existir menos informação no QUALIS.

Aqui detetou-se que este desvio nos históricos corresponde aos dados que foram eliminados do *Controlab*. Ou seja, quando são gerados relatórios e/ou ensaios em cada “refresh” são lidos e armazenados na DW. Uma vez lidos os dados não eram eliminados deste, isto influenciava as análises que eram efetuadas via QUALIS, tornavam-se distintas das efetuadas com recurso ao módulo “Estatísticas” disponível no *Controlab*.

Como apresentado em Dill (2002), a característica não-volátil tal como foi apresentada por Dr. Inmon já não é tão frequente nas atuais DW. Neste sentido, o QUALIS foi adaptado para eliminação dos boletins que não estão presentes nos históricos dos *softwares*. Assim os dados que são lidos na altura do processo ETL são armazenados no DW, caso estes tenham sido eliminados posteriormente, atualmente não são visíveis para o utilizador e são ignorados nos cálculos efetuados pelo cubo.

3.1.5 Alteração de ensaios *Controlab*

Para que fosse possível extrair a informação presente nas fontes de dados para o QUALIS teve de proceder-se a ajustes nos *softwares* fontes.

O ensaio Visual é um ensaio laboratorial bastante específico. Este permite-nos através de uma amostragem aferir qual é a classe do lote, tal como a sua percentagem de defeitos.

¹¹O *Controlab* e *SPC* são *softwares* que utilizam um ambiente de multi-utilizadores que com um simples clique a informação fica atualizada. Como relatado em 2.1 o OLAP não atualiza ao momento, o QUALIS apenas faz o “refresh” dos dados duas vezes por dia. A quantidade de dados lida pelo cubo é enorme e a diversidade de transformações que são necessárias no ETL, implica que cada atualização demore algum tempo, tornando-se inviável a constante atualização.

Devida às propriedades da cortiça é difícil uma classe ser pura: as classes em geral contêm percentagens da classe superior e classe inferior. Existem 8 classes puras: Flor, Extra, Superior, 1ª, 2ª, 3ª, 4ª e 5ª (rolha Natural). Neste ensaio as rolhas são desdobradas nas várias classes, e as que têm defeitos¹² são separadas e catalogadas segundo a gravidade do defeito a que pertencem (crítico, maior ou menor).

Na maior parte dos casos, o resultado da desdobra é comparada com a amostra padrão da AI ou com a amostra acordada com o cliente final. Nestes, o cliente e a empresa acordam quais são as percentagens de rolhas que aceitam fora da classe pura. Tendo em conta estas referências é efetuado o cálculo do desvio à classe, caso seja positivo indica que o lote preparado é de classe superior à amostra padrão.

Tal como os restantes ensaios, pode existir a repetição deste teste. Enquanto que nos restantes ensaios o último ensaio é o considerado para a validação final (único impresso nos relatórios), na classe Visual pode ser considerado meramente o último teste ou os vários testes efetuados na validação do lote. Isto é, cada teste tem associado o desdobramento da classe, desvio e defeitos. Quando existe repetição dos ensaios fica-se com a amostra global para as classes que resulta da compilação das várias amostragens e o desvio à classe pode ou não ser influenciado por todos os testes efetuados. Inicialmente, a classe visual tinha dois campos para efetuar estes cálculos: “Completo” e “Considerado em Resumo”.

A opção “Considerado em Resumo” indica se o desvio do ensaio é considerado ou não para o desvio global do teste. Os ensaios que continham esta opção ativada, não eram considerados no QUALIS, o que não fazia sentido pois essa indicação é dada pela não ativação da *flag* “Estatística” a nível do ensaio.

Desta forma, é necessário lidar com esta temática para que seja possível a exportação de todos os ensaios e para que sejam devidamente identificados. Em suma, o ensaio visual pode ter três opções:

1. estatística e resumo: o teste conta para efeitos estatísticos e o seu desvio influenciará o desvio global da classe;
2. resumo: não deverá ser contabilizado nas análises estatísticas, mas será contabilizado no cálculo do desvio;
3. estatística: o resultado final não influenciará o desvio final, mas é contabilizado como carga de trabalho.

Assim, com a aplicação desta transformação nos boletins de *Controlab* ficará a anotação de qual foi realmente a finalidade do teste. Foi introduzida esta implementação a nível do *Controlab*, sendo por defeito assumido que o teste será para fins estatísticos e para resumo. Desta forma, a *flag* de “Estatística” fica igual em todos os ensaios. No QUALIS teve de ser gerado um novo campo denominado “QualisStatus”

¹²Os defeitos podem resultar do processo produtivo (mal aglomeradas, mal marcadas) ou de causas externas ao processo (bicho, ano seco).

onde está identificada qual foi a opção ativada no *Controlab*, que permite estudar a classe visual nos vários níveis.

O ensaio de forças de extração teve de ser adaptado no *software*, porque este ensaio tem associado um espaço temporal.

A força de extração indica-nos qual a força necessária para conseguir extrair a rolha da garrafa. Em laboratório este tipo de forças são quantificadas através de equipamentos apropriados como o ExtraLab, sendo expressas em decaNewton (daN).

As normas especificam que estas devem situar-se entre os 20 e 40 daN para um bom desempenho da rolha na garrafa. Valores elevados de força de extração indicam uma difícil extração da rolha e valores baixos podem indiciar uma deficiência funcional. Os problemas de forças de extração estão na sua maioria associados ao tratamento de superfície da rolha, perfis de garrafa VS diâmetro da rolha ou má inserção das rolhas nas garrafas¹³.

As forças de extração não são inertes ao longo do tempo, têm tendência a diminuir (APCOR, 2011), ou aumentar dependendo do tipo de tratamento. Segundo o Código Internacional das Práticas Rolheiras (CELIEGE, 2011), as empresas devem seguir a norma “ISO 9727 – 5 Rolhas Cilíndricas - Ensaaios físicos - Parte 5: Determinação da força de extração” para o controlo destas. Esta ISO indica que o ensaio de forças de extração deve ser efetuado após 24 horas de engarrafamento.

Este ensaio pode ter vários ensaios sendo em espaços temporais iguais e/ou distantes. O *ControLab* permitia o registo de vários ensaios para este teste, no entanto não podiam estar associados ao mesmo espaço temporal. Quando era necessário efetuar a repetição, o *software* obrigava a indicar um tempo diferente do real, não existindo um histórico coerente.

Portanto, foi sugerido que *software* fosse ajustado para ser possível introduzir vários ensaios com o mesmo tempo de extração, onde existisse uma tabela onde são pré-definidos os possíveis tempos de extração da rolha. Quando fosse adicionado um novo ensaio, contempla-se a opção de escolha dos espaços temporais estabelecidos nesta tabela, independentemente dos ensaios anteriormente realizados.

Dada a configuração inicial deste ensaio, a informação que é extraída pelo QUALIS somente diz qual foi a ordem de realização do ensaio, em nada indicando o seu tempo em garrafa na altura da extração. Para tal, foi introduzido um campo denominado “Extraction Time”, que permite fazer análises comparativas dos resultados obtidos sem ocorrência de mistura de espaços temporais, que falseava os resultados.

¹³A Amorim & Irmãos, S.A. apoia os seus clientes com a informação das boas práticas de engarrafamento para um bom desempenho das suas rolhas, pois alguns problemas que surgem não se devem às rolhas mas a problemas de engarrafamento, armazenamento e transporte dos seus vinhos. Um apoio prestado fundamental é a preconização dos calibres e tipo de rolhas adequados para as suas garrafas VS vinho. As garrafas têm diferentes dimensões de gargalos em que uma mesma rolha tem um comportamento dispar nas várias garrafas.

3.1.6 Alteração ao *SPC*

O *software SPC* como o próprio nome sugere é utilizado para a monitorização e controlo estatístico do processo produtivo, onde são registados parâmetros do produto e/ou processo, nos vários postos de controlo, permitindo a análise estatística dos dados através de cartas de controlo de Shewhart.

Cada máquina/posto de controlo tem definido nas suas “Configurações de Registo” qual a carta de controlo utilizada para a monitorização do processo. Nestas configurações estão definidos os campos de identificação de registo, características controláveis, tal como a dimensão da amostra (por defeito) para cada posto de inspeção. As estatísticas descritivas que dependem da dimensão da amostra (por exemplo: desvio-padrão), no *software* os seus cálculos são efetuados com base nesta dimensão da amostra por defeito, o que pode levar à existência de erros nos cálculos. Caso seja validado um novo plano de inspeção que contemple uma dimensão do subgrupo diferente é necessário registar esta nova dimensão nas configurações, para que os novos registos conttenham a dimensão correta. Porém, analisando os registos anteriores à data de alteração da dimensão do subgrupo, o *software* utilizava nos cálculos a atual dimensão definida por defeito, e não reconhece a efetiva dimensão dos subgrupos em estudo.

Para lidar com esta situação, na carta de controlo em estudo, caso esteja contemplado pelo menos um registo em que a dimensão efetiva seja diferente à dimensão registada por defeito nas configurações é disputado um alerta com essa indicação. Assim, o utilizador prosseguindo com a análise fica alertado que é uma análise incorreta.

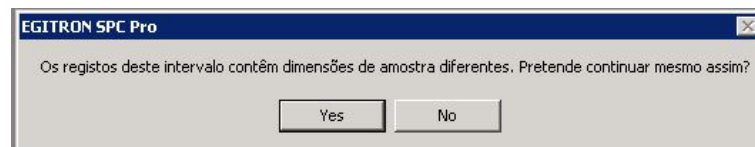
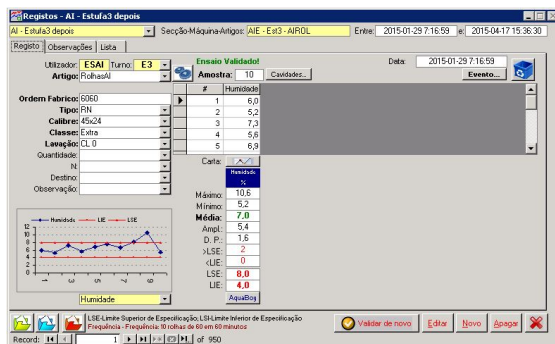
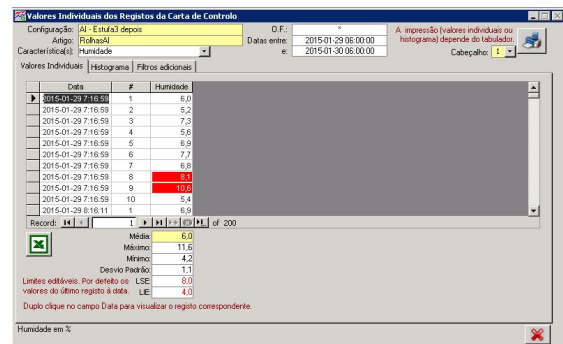


Figura 3.5: Alerta introduzido no *SPC* para subgrupos de dimensões diferentes

O *software SPC* é limitador a nível de um controlo de processo numa fase primária do estudo/investigação, pois não permite uma análise completa em termos de valores individuais. Os dados individuais que são registados no *SPC* apenas são usados para o cálculo dos valores médios, amplitudes e desvio-padrão dos subgrupos, existindo uma possibilidade de visualização das estatísticas descritivas entre os valores individuais. Depois é com as estatísticas descritivas dos subgrupos que são efetuados todos os cálculos para o controlo estatístico do processo e para construir as cartas de controlo.



(a) Registo de um subgrupo



(b) Listagem valores individuais dos subgrupos

Figura 3.6: Visualizações dos cálculos efetuados com valores individuais no *SPC*

Figura 3.6a é a janela obtida no *software* para quando um ensaio (subgrupo) é validado. Aqui temos acesso aos valores individuais e as suas estatísticas descritivas, bem como toda a informação associada a este registo. As amostragens coletadas nas datas compreendidas na seleção são resumidas numa tabela denominada “Registos de Cartas de Controllo” (Figura 3.7).

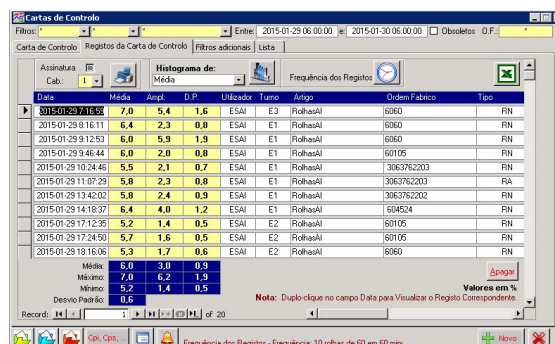


Figura 3.7: Listagem dos subgrupos Carta de Controllo

No final desta listagem estão calculadas as estatísticas descritivas das estatísticas descritivas dos subgrupos, sendo estas as apresentadas nas Cartas de Controllo e usadas para o cálculo dos índices de capacidade do processo.

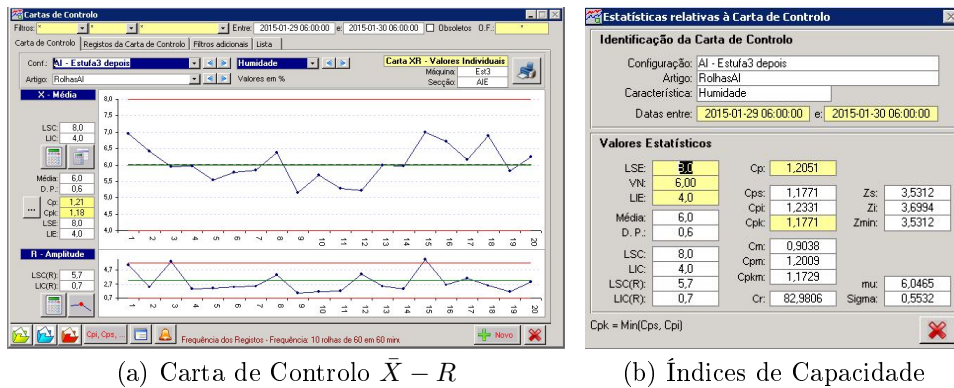


Figura 3.8: Carta de Controle $\bar{X} - R$ e respectivos índices de capacidade, ROSA EVO 3

Neste caso, a análise corresponde a um dia de trabalho da estufa ROSA EVO 3 da AI, onde é inspecionada a humidade da rolha, obtém-se um desvio-padrão de 1.1208 entre valores individuais de todas os subgrupos e um desvio-padrão de 0.5532 entre médias dos subgrupos. Esta disparidade é explicada pela menor oscilação entre valores médios comparativamente com os valores individuais, ou seja, apesar de existirem valores individuais bastante distantes em termos médios estas oscilações são encobertas.

Um parâmetro fundamental para o CEP é a variabilidade inerente ao processo produtivo, que pode ser traduzida pelo desvio-padrão. Assim, estudar o processo por valores individuais ou por valores médios, pode obter-se resultados/conclusões opostas.

Analisando os índices de capacidade mais usuais para os dados anteriores, onde os limites de especificação são $LES = 8\%$ e $LEI = 4\%$, obtém-se:

	Valores individuais subgrupo	Valores médios subgrupo
Desvio-padrão	1.1208	0.5532
Média	6.0465	6.0465
C_p	0.5948	1.2051
C_{pk}	0.5810	1.1771

Tabela 3.2: Índices de Capacidade para valores individuais e médios dos subgrupos

No presente exemplo, o desvio-padrão do subgrupo calculado com todos os valores individuais dos subgrupos corresponde aproximadamente ao dobro do desvio-padrão calculados através dos valores médios dos subgrupos, o que traduz um processo produtivo razoável (valores médios) em confronto com um processo produtivo incapaz (valores individuais).

Este cálculo das estatísticas descritivas sem recorrer diretamente aos valores individuais dos subgrupos pode ser problemático (como ilustrado anteriormente e que será explicado com maior pormenor no caso de estudo das Estufas ROSA EVO (secção 3.2), ver página 54).

Globalmente a Amorim & Irmãos, S.A. ainda está numa fase embrionária a nível do CEP. Para que exista uma maior clareza do comportamento do processo produtivo foi criado um ficheiro em Excel¹⁴. Este foi programado para que quando são inseridos os valores individuais recolhidos na inspeção do processo produtivo, os cálculos sejam efetuados com base nestes.

Neste ficheiro os índices de capacidade são calculados tendo em conta a estimação pela amplitude (Equação 2.16) e pelo desvio-padrão (Equação 2.17), tal como pelo desvio-padrão real dos valores individuais (Equação 2.8).

Para o dia de trabalho ilustrado anteriormente, obteve-se os seguintes resultados:

$$\hat{\sigma}_0 = \frac{\bar{R}}{d_2} = \frac{2.9600}{3.0780} = 0.9617. \quad (3.1)$$

$$\hat{\sigma}_0 = \frac{\bar{S}}{c_4} = \frac{0.9486}{0.9727} = 0.9752. \quad (3.2)$$

$$s = \sqrt{\frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n - 1}} = 1.1208. \quad (3.3)$$

Pelas estimativas do desvio-padrão obteve-se valores inferiores ao desvio-padrão real dos valores individuais. Assim, neste caso os índices de capacidade mais usuais são dados por:

Estimado por \bar{R} barra		Estimado por \bar{s} barra		Estimado por DP real	
Cp	0,69324	Cp	0,68362	Pp	0,59480
Cpks	0,67713	Cpks	0,66772	Ppks	0,58097
Cpki	0,70936	Cpki	0,69951	Ppki	0,60862
Cpk	0,67713	Cpk	0,66772	Ppk	0,58097

(a) Índices de capacidade por \bar{R} (b) Índices de capacidade por \bar{S} (c) Índices de Capacidade por desvio-padrão real

Figura 3.9: Índices de capacidade mais usuais

Neste ponto da análise estamos perante vários valores para C_p e C_{pk} . O “analista” tem uma panóplia de indicadores do seu processo produtivo que o ajudará nas tomadas de decisão em prol da eficácia do processo.

¹⁴Foi escolhido este programa, por ter sido em conta a possibilidade de extração dos valores individuais registados no *SPC* com recurso ao QUALIS.

Este ficheiro contempla uma secção que apresenta o histograma¹⁵ dos valores individuais para que seja possível aferir sobre a Normalidade dos valores adquiridos.

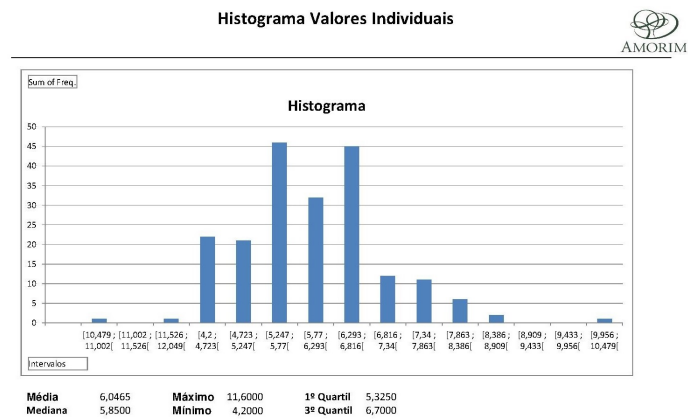


Figura 3.10: Histograma Valores individuais

Neste ainda está disponível uma representação gráfica de todos os valores individuais recolhidos no período de tempo escolhido em simultâneo com os limites de especificação. Através deste consegue-se concluir a existência/ausência de rolhas não conformes.

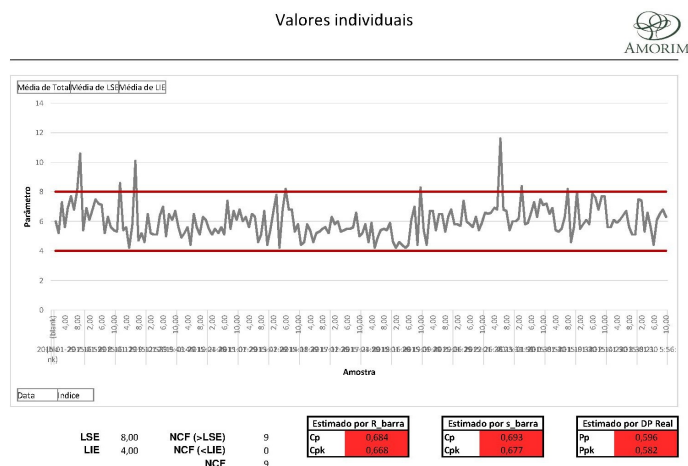


Figura 3.11: Especificações para os valores individuais

De forma análoga ao *software SPC*, foram programadas as Cartas de Controlo

¹⁵O número de classes deste é dado por: \sqrt{n} .

$\bar{X} - R$ e $\bar{X} - S$, com os respetivos índices de capacidade mais usuais e estimação dos limites de controlo pelos fatores tabelados que dependem da dimensão dos subgrupos recolhidos¹⁶. Os cálculos são efetuados segundo os limites de controlo apresentados na página 23.

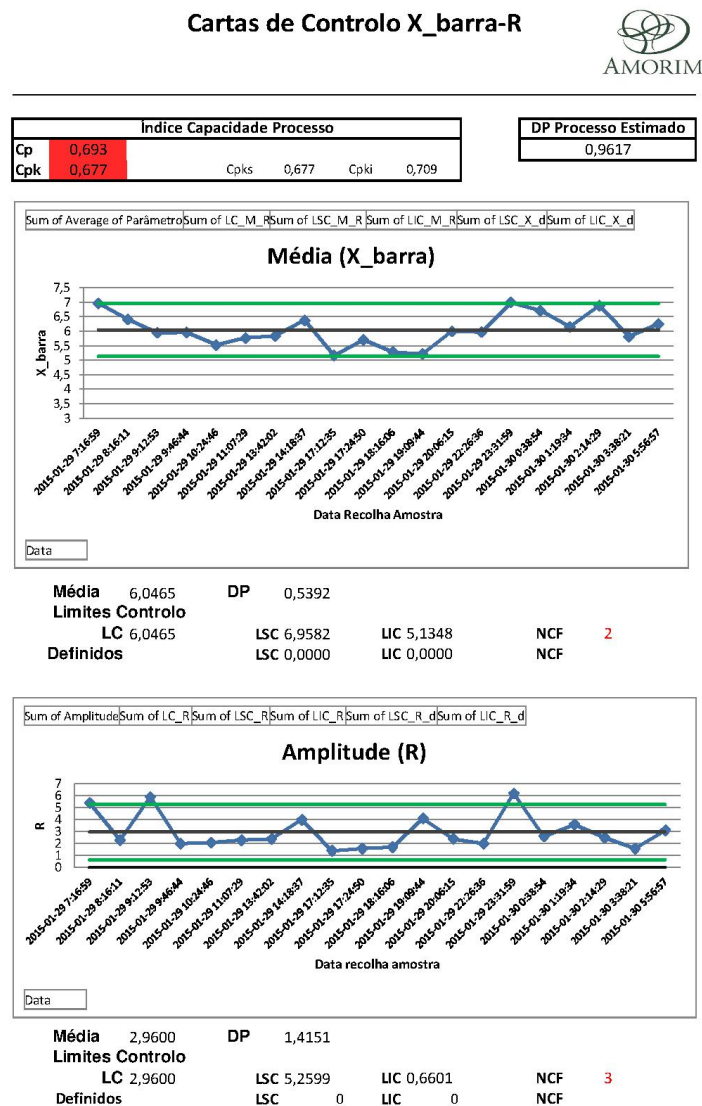


Figura 3.12: Carta de Controlo $\bar{X} - R$

¹⁶Neste ficheiro os cálculos são efetuados pela análise dos valores individuais dos subgrupos. Este ficheiro não permite a análise conjunta de registos com dimensões heterogêneas.

Cartas de Controlo \bar{X} -s

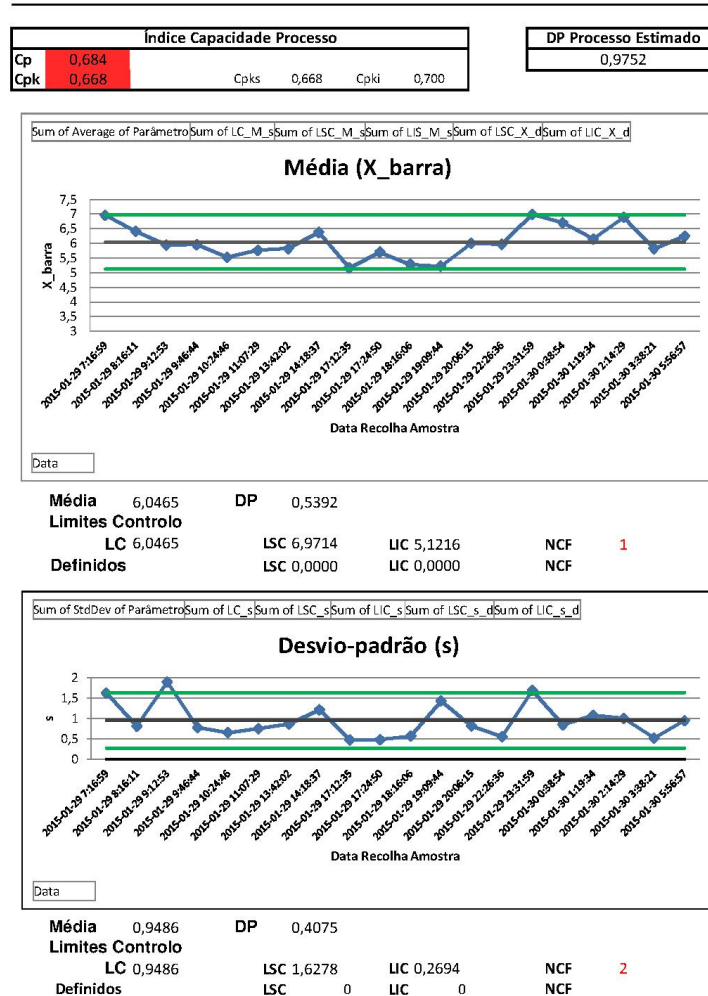


Figura 3.13: Carta de Controlo $\bar{X} - s$

3.1.7 Registos do *SPC* no QUALIS

No início do estágio curricular a informação do *software SPC* não era extraída eficazmente para o cubo, o que não permitia efetuar as consultas pretendidas.

Para efetuar uma consulta de registos no *software* é necessário ter a informação

da configuração¹⁷ em que foi registada, o dia e hora do registo. Estes campos não estavam disponível no cubo.

No cubo a “chave primária” de cada registo era o “Report Number”. Este é um código numérico único, atribuído a cada registo pelo *SPC* no momento de gravação na base de dados. Este número fica oculto para o utilizador, mas é de extrema importância para que não ocorra erros no cruzamento das tabelas de informação do SI. Assim, o utilizador não conseguia obter a rastreabilidade entre os dois SI’s. No cubo existiam alguns campos relativos aos relatórios de *SPC* (Turno, Secção, Grupo de Artigos, Data Injeção, Máquina), e ainda assim não eram suficientes para dar resposta ao pretendido.

O *software* para cada configuração permite utilizar 15 campos para identificação dos registos (para além dos campos obrigatórios: Data, Operador, Turno, Ordem Fabrico e Artigo) e outros 9 campos para os possíveis parâmetros de controlo que originam as Cartas de Controlo. O QUALIS para identificar estes campos utilizava a nomenclatura presente no *SPC*. Caso algumas configurações contivessem a mesma informação não era reconhecida pelo cubo como igual.

Neste projeto foram identificados todos os campos que seriam necessários exportar para o cubo, e quais poderiam ser ajustados aos já existentes. Esta opção de ajustamento dos campos foi fulcral para manter um cubo intuitivo e simples, que facilita-se as consultas. No sentido de uniformizar estas identificações no cubo, foi efetuado um *match* entre as configurações para a identificação de registo e parâmetros de controlo, das várias UI’s. Desta forma, na fase de processamento de ETL estes passaram a ser identificados de igual forma em postos de controlo distintos.

Atualmente, já é possível analisar os registos de *SPC*, através do QUALIS.

3.2 Controlo Estatístico do Processo - Caso de Estudo

Foi efetuado um estudo¹⁸ do CEP nas Estufas ROSA EVO da AI. Esta secção da fábrica contém 4 estufas, onde é efetuado o tratamento ROSA EVO às rolhas naturais após serem lavadas. Este tratamento ocorre já numa fase de finalização do processo produtivo.

Cada estufa tem programas a que as rolhas são submetidas por forma a que haja uma redução da humidade e de TCA. Devido à extensão das estufas, as rolhas podem estar dentro destas durante horas. As rolhas são colocadas nas estufas por OF¹⁹, com intervalos de 30 minutos para que não ocorra mistura de OF’s.

¹⁷A configuração é a identificação específica do PC. Um PC pode conter mais que uma configuração.

¹⁸Ao longo deste trabalho será utilizado o *software* gratuito R, versão 3.1.2..

¹⁹No entanto, quando a quantidade de rolhas por OF é grande a OF é repartida, para que não ocorra uma perda de eficiência do tratamento ROSA EVO.

3.2.1 Amostragem

Neste PC está definido no plano de inspeção e ensaio efetuar-se uma amostragem de tipo FSI, onde são recolhidas 10 rolhas de hora a hora²⁰, para inspeção de humidade. Devido às características da cortiça existe uma heterogeneidade de humidade entre lotes, daí ser importante um controlo deste parâmetro no processo produtivo. Este controlo de humidade das rolhas é efetuado a todas as OF's, antes e após passarem pelo ROSA EVO.

O controlo de humidade antes estufa é efetuado uma vez por OF, tendo como finalidade decidir em que estufa a alocar. Após estufa, esta inspeção é importante para avaliar o teor de redução de humidade, e para verificar se a OF está conforme. Para cada OF é retirada uma amostragem quando esta começa a sair da estufa, caso tenha uma grande quantidade de rolhas, de hora a hora é recolhida nova amostragem.

Para cada controlo é efetuado um registo no *SPC* na respetiva configuração (antes ou após), com a identificação de registo (data, hora, turno, OF, artigo, tipo, calibre, classe e lavação da rolha) e os resultados de humidade percentual das rolhas obtidos com recurso ao Aquaboy.

3.2.2 Cartas de Controlo de Shewhart

A fábrica não é climatizada, assim as rolhas estabilizam em ambiente natural. Por esta razão existe a necessidade de ter os limites de especificação em função da estação do ano:

- Primavera/Verão (Entre a semana 13 e a 40): 5% a 9%
- Outono/Inverno (Restantes semanas): 4% a 8%

Aqui serão tratados os dados recolhidos após as rolhas terem sido submetidas ao tratamento ROSA EVO, para as quatro estufas. Os dados referentes à estufa 1 serão identificados com d1, à estufa 2 por d2 e assim sucessivamente. Este estudo terá em conta a estação do ano.

Neste PC são usadas as cartas de controlo de Shewhart $\bar{X} - R$. Para fazer o estudo dos limites de controlo para estas cartas, a distribuição da característica do processo em estudo tem de seguir uma distribuição Normal (ou aproximadamente), como referido no capítulo anterior. Neste PC não existe um conhecimento à partida da média e desvio-padrão do processo produtivo, sendo necessário estimá-los.

Efetuuou-se o estudo do **comportamento da humidade** tendo em conta o histórico dos valores individuais de 2014, para avaliar a condição da Normalidade dos dados. Pelos seguintes histogramas, é óbvio que todas as estufas apresentam um enviesamento positivo dos valores individuais.

²⁰Dado o intervalo de tempo entre entradas de OF's nas estufas, pode não ser possível efetuar a inspeção de de hora em hora.

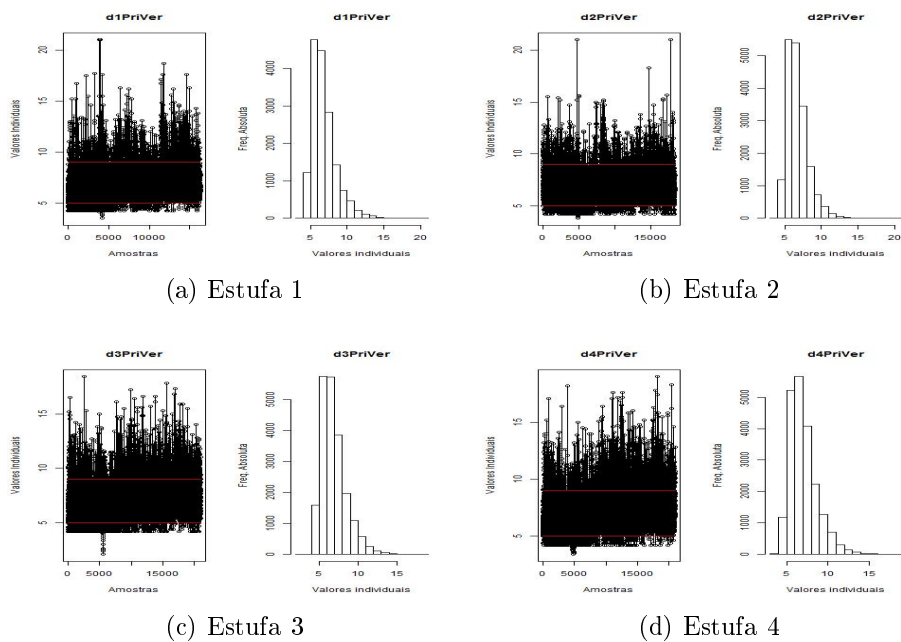


Figura 3.14: Histogramas de dados Primavera/Verão de 2014

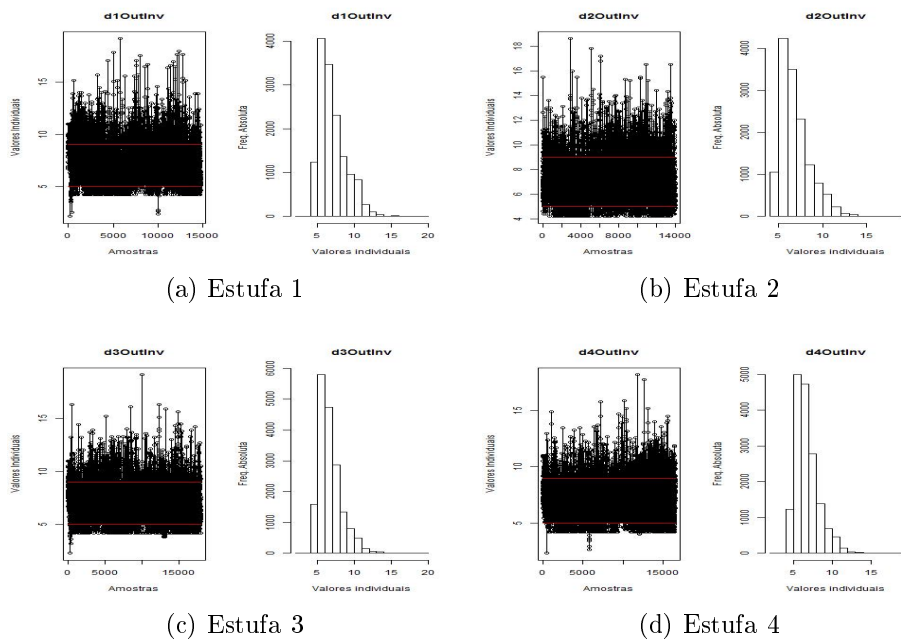


Figura 3.15: Histogramas de dados Outono/Inverno de 2014

Para validar as conclusões anteriores calculou-se as seguintes medidas²¹:

Estufa/Estação	Média	Desvio-padrão	Coef. de Assimetria	Coef. de Achatamento
d1PriVer	6.8415	1.6911	1.4019	6.4035
d1OutInv	7.0382	1.8936	1.0615	4.2857
d2PriVer	6.7352	1.4727	1.2729	6.4251
d2OutInv	6.8852	1.7246	1.1565	4.6837
d3PriVer	6.8685	1.6523	1.1705	5.2974
d3OutInv	6.6665	1.5814	1.2114	5.0766
d4PriVer	7.0625	1.7165	1.2025	5.4484
d4OutInv	6.7238	1.5327	1.1481	5.0289

Tabela 3.3: Média, Desvio-padrão, coeficiente de assimetria e coeficiente de achatamento, dos registos de inspeção de 2014

Os coeficientes de assimetria são ligeiramente superiores a 1, em todas as estufas e em ambas as estações, o que traduz o enviesamento positivo dos dados. Quanto aos coeficientes de achatamento são bastantes elevados. Conclui-se que a distribuição Normal tem um achatamento inferior em relação aos dados recolhidos nas várias estufas em 2014.

Estas constatações não corroboram para a aceitação da Normalidade das distribuições de percentagens de humidade das rolhas após passarem nas estufas de tratamento ROSA EVO.

Estufa/Estação	Mínimo	1º Quartil	2º Quartil	3º Quartil	Máximo	Amplitude
d1PriVer	3.5	5.6	6.5	7.7	21.0	17.5
d1OutInv	2.1	5.6	6.6	8.0	19.2	17.1
d2PriVer	3.8	5.6	6.5	7.5	21.0	17.2
d2OutInv	4.2	5.6	6.5	7.7	18.6	14.4
d3PriVer	2.1	5.6	6.6	7.7	18.4	16.3
d3OutInv	2.3	5.5	6.3	7.4	19.1	16.8
d4PriVer	3.4	5.8	6.8	7.9	19.0	15.6
d4OutInv	2.3	5.6	6.5	7.4	18.2	15.9

Tabela 3.4: Medidas de dispersão, dos registos de inspeção de 2014

Para confirmar as hipóteses anteriores, realizou-se um teste de Normalidade aos dados recolhidos. Aplicando o teste de Kolmogorov-Smirnov aos dados da humidade percentual após estufa, para o ano 2014, foram obtidos os seguintes resultados:

²¹Para o cálculo dos coeficiente de assimetria e achatamento foi usado o *package* 'moments' do R, consultar Komsta (2015).

Estufa/Estação	Estatística de teste	<i>p-value</i>
d1PriVer	0.1031	$< 2.2 \times 10^{-16}$
d1OutInv	0.1048	$< 2.2 \times 10^{-16}$
d2PriVer	0.0877	$< 2.2 \times 10^{-16}$
d2OutInv	0.1071	$< 2.2 \times 10^{-16}$
d3PriVer	0.0904	$< 2.2 \times 10^{-16}$
d3OutInv	0.1000	$< 2.2 \times 10^{-16}$
d4PriVer	0.0946	$< 2.2 \times 10^{-16}$
d4OutInv	0.0967	$< 2.2 \times 10^{-16}$

Tabela 3.5: Resultado do teste Kolmogorov-Smirnov dos registos de inspeção de 2014

Para um nível de significância de 5%, é rejeitada a hipótese nula do teste de Normalidade de Kolmogorov-Smirnov, devido os baixos valores do *p-value*. Este *p-value* indica que se rejeita H_0 , podendo ocorrer a hipótese alternativa em todas as estações do ano e em todas as estufas.

Esta não Normalidade associada às estufas poderá ser explicada pela diversidade de **tipologia de rolhas** ou pela **lavação** das rolhas. O principal objetivo da lavação é garantir a desinfecção da rolha. Não obstante existem várias lavações, por exemplo: Clean 0, Clean C, Clean 2000, que podem conferir cores diferentes às rolhas. Assim, foram analisadas as distribuições dos valores individuais das amostragens para cada estufa, tendo como critério o tipo de rolha ou a lavação destas. Mesmo assim, a percentagem de humidade nas rolhas não segue uma distribuição Normal, nem por tipo de rolha, nem por lavação.

Outra hipótese para esta não Normalidade poderá estar associada à **amostragem realizada**. Com o auxílio da equipa da AI foi efetuado um estudo de amostragens neste PC.

3.2.2.1 Estudo amostragens no ROSA EVO

Para este estudo foram acompanhados lotes de rolhas Naturais com lavação Clean 2000, durante uma semana.

Este estudo teve o objetivo de verificar qual o efeito de retirar as amostragens logo após a chegada dos contentores da lavação em confronto com a amostragem antes de entrar na estufa; amostragem mais superficial (concentrada ou dispersa) em confronto com amostragem superficial e profunda no contentor; amostragem após lavação ou antes entrada na estufa; a recolha de amostra após passagem na estufa concentrada em confronto com uma amostragem dispersa ao longo da largura do tapete. Para tal, foi usada a seguinte metodologia para a recolha de amostras:

- Antes Estufa: À saída da lavação, recolha de 10 rolhas de vários pontos do

contentor. À entrada da estufa, recolha de 5 rolhas do cimo e 5 rolhas do meio do contentor.

- Após Estufa: Recolha no centro e extremidades do tapete (4 rolhas em cada).

As amostragens recolhidas pela metodologia proposta foram registadas num “Diário de Bordo” criado para o efeito. Após a lavação foram controlados 42 lotes, apenas 39 lotes passaram em todas as fases do estudo. Os lotes podem ter várias dimensões, não obstante foi apenas analisado o primeiro contentor (capacidade aproximada de 2 ML de rolhas), para garantir que ao longo do percurso seria sempre esse o contentor controlado. Os lotes inspecionados corresponderam a 4 calibres e 16 classes distintas. Tendo em conta os dados recolhidos foram efetuadas as seguintes análises:

- **Antes Estufa**

1. **Amostragem Superficial Concentrada VS Dispersa após Lavação**

Esta análise confronta a amostragem atual (superficial com uma recolha concentrada) com a recolha superficial da proposta (mesma quantidade de rolhas em vários pontos no cimo do contentor).

Não é expectável que haja uma diferença significativa, dado que as rolhas são lavadas e num curto espaço de tempo são inspecionadas.

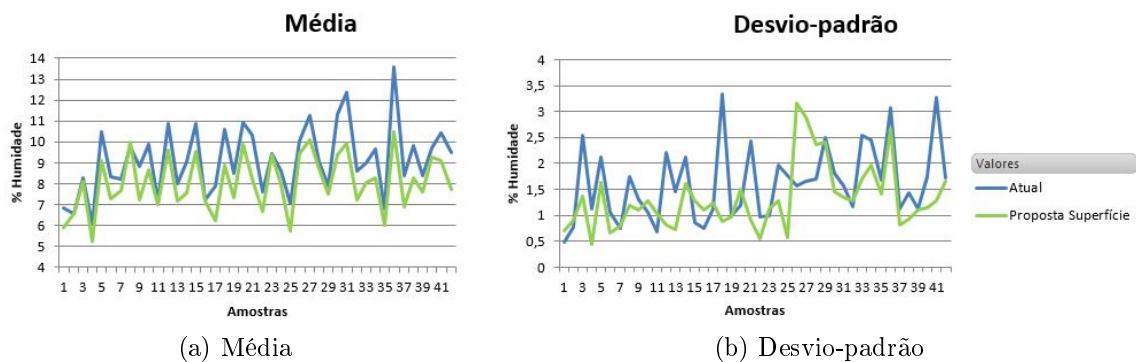


Figura 3.16: Amostragem Superficial após Lavação

Com a atual recolha de amostras obteve-se humidades superiores em quase todos os lotes controlados, em média 1%. Os valores individuais numa recolha atual são mais heterogéneos em comparação com uma amostra dispersa. Os diferentes modos de recolha podem influenciar na decisão da ação a tomar com os lotes.

2. Amostragem Superficial VS Profundidade após Lavação

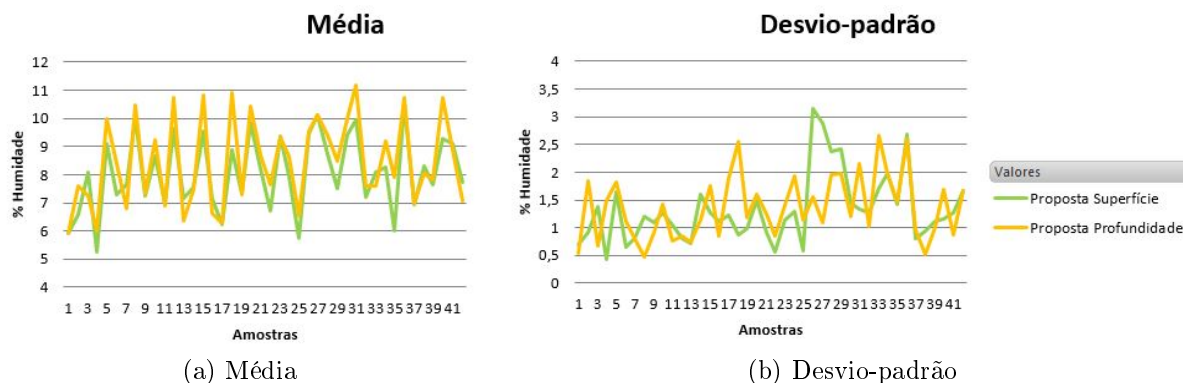


Figura 3.17: Amostragem Superficial VS Profundidade após Lavação

Estas amostragens foram ambas retiradas de vários pontos do contentor, nos níveis especificados. Tendo em conta a distribuição das humidades no contentor, em termos médios não é notória uma grande diferença entre as amostras superficiais e amostras mais profundas, verificando-se maiores amplitudes na camada superficial. No entanto, como expectável, na camada superficial é detetada uma percentagem de humidade ligeiramente mais reduzida devido à maior área de exposição com o ar.

3. Amostragem Atual VS Proposta após Lavação

Na confrontação da amostragem superficial concentrada do lote com as suas amostragens superficial e profundidade, os resultados são diferentes. Pelos pontos anteriores, conclui-se que nesta fase será indiferente o nível do contentor para a recolha da amostra de inspeção; por sua vez nas amostras concentradas ou dispersas já são visíveis algumas diferenças. Na inspeção após lavação deve ser ponderada a continuação de recolha concentrada VS a recolha dispersa.

Nas amostragens atuais observam-se humidades superiores, sendo as maiores amplitudes registadas na amostragem proposta (vários níveis) no momento de chegada dos contentores após lavação.

4. Amostragem Após Lavação VS Antes Entrada Estufa

Os contentores quando chegam às estufas, após inspeção da humidade, são acondicionados nos corredores entre as estufas e por vezes por questões de espaço têm de ser estacionados afastados das estufas. A secção das estufas não é climatizada, assim as temperaturas são mais elevadas entre estufas do que no estacionamento. 83% dos lotes aguardaram a entrada na estufa no corredor entre estufas, e os restantes ficaram afastados das estufas.

Nesta fase não foi possível controlar 2 dos lotes iniciais, assim serão analisados 40 lotes nesta fase. Na amostragem “Atual” foi considerada a amostragem que o operador efetuou à chegada da lavação e na “Proposta Superficial” foi a amostra recolhida na superfície do contentor no momento imediatamente antes da entrada na estufa.

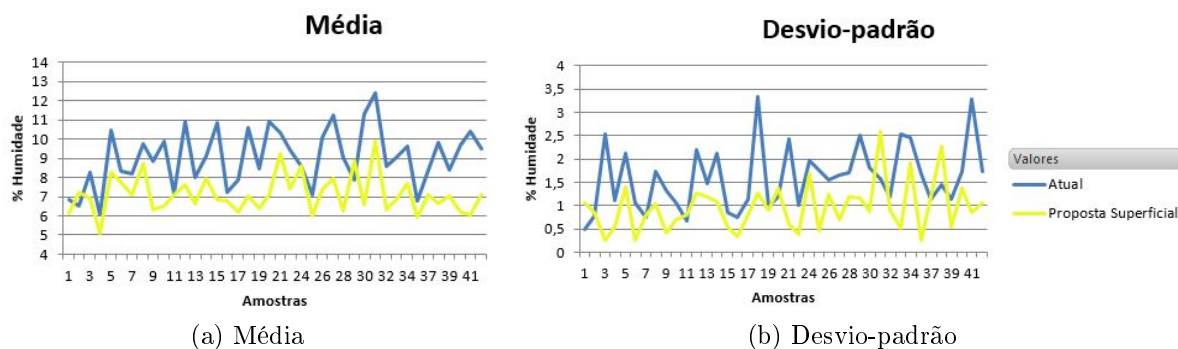


Figura 3.18: Amostragem Atual VS Proposta após Lavação

Os lotes inspecionados na entrada da estufa têm humidades dispares às obtidas após lavação. A média das percentagens de humidade é superior quando o lote é controlado à chegada ao setor, registando uma maior variabilidade entre os valores individuais.

Também foi avaliado o comportamento das percentagens de humidade em dois níveis do contentor, no momento de entrada na estufa.

É evidente a discrepância entre recolhas no cimo e fundo do contentor, neste PC. Nas recolhas do fundo dos contentores revelam valores de humidade superiores, chegando a atingir uma diferença máxima de 4.76% entre os dois “níveis”.

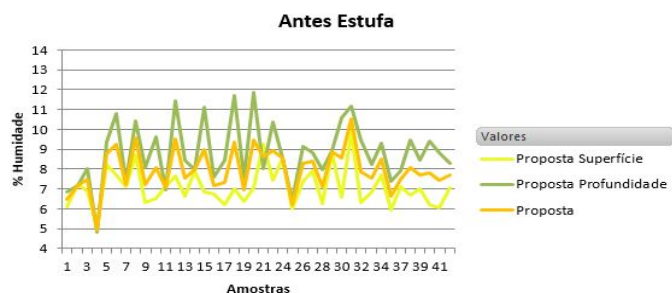


Figura 3.19: Média das amostragens por lote

Pela Figura 3.19 é de esperar que não se obtenha qualquer semelhança entre a metodologia proposta e a efetuada atualmente. Para entender se a

alocação dos lotes têm influência nesta proposta de controlo, verificou-se as diferenças entre os níveis do contentor. No entanto, existe uma grande discrepância entre os lotes em estacionamento e corredores.

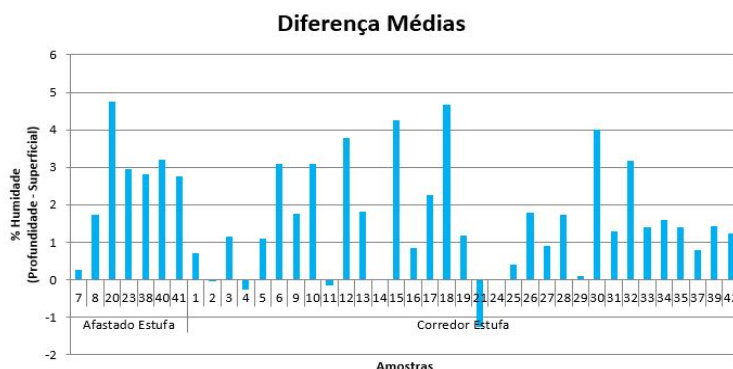


Figura 3.20: Diferença de humidade entre fundo e cimo contentor, antes estufa

Os lotes que inicialmente são alocadas no estacionamento ilustram uma notória diferença de humidade entre as rolhas superficiais e mais profundas no contentor. Os lotes estacionados entre estufas têm uma diferença na grandeza de 1.49%, enquanto outros perfazem um valor de 2.63%.

5. Amostragem Atual VS Proposta (superfície e profundidade)

Quando efetuadas amostragens com recolha de rolhas à superfície e profundidade do contentor, obtêm-se percentagens de humidades inferiores e com maior variabilidade entre os seus valores individuais no controlo Antes Estufa.

• Depois Estufa

Nesta fase de inspeção da humidade não foi possível acompanhar 3 lotes iniciais, assim esta secção apenas considera 39 lotes. Devido ao fluxo de trabalho das estufas ROSA EVO não foi possível passar todos os lotes na mesma estufa, a distribuição dos lotes pelas várias estufas foi:

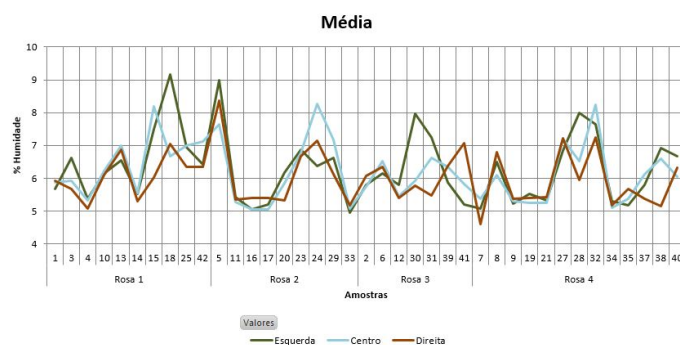
Estufa ROSA EVO	1	2	3	4
Nº Lotes	10	9	7	13

Tabela 3.6: Distribuição lotes pela Estufas ROSA EVO

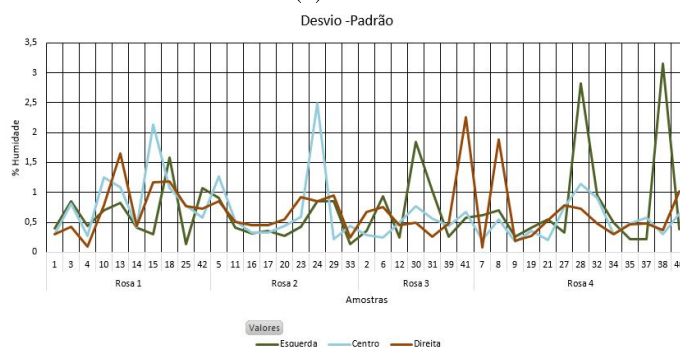
Na semana do estudo os limites de especificação eram 4% e 8%.

1. Amostragem distribuição no tapete

Devido à pouca quantidade de lotes inspecionados e por não passarem na mesma estufa, não é possível retirar conclusões relativamente a cada posição específica de saída do tapete.



(a) Média



(b) Desvio-padrão

Figura 3.21: Amostragens distribuição pelo tapete

2. Amostragem Atual (concentrada) VS Proposta (dispersa)

Pelas amostragens atuais todos os lotes têm a sua média conforme. Enquanto, que pela metodologia proposta (mistura dos três pontos do tapete) um lote tem a sua média fora do LES e outra fora do LEI. As amostragens propostas geralmente tem uma menor dispersão.

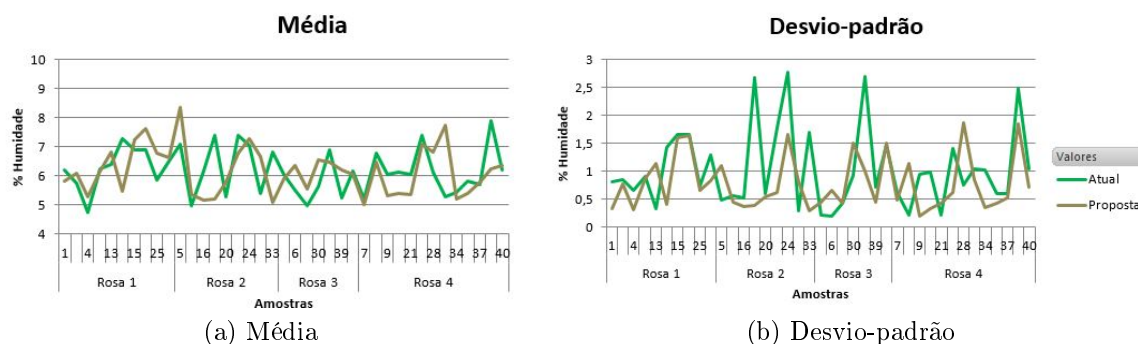


Figura 3.22: Amostragem Atual VS Proposta após Estufa

Este estudo não foi conclusivo para a explicação da não Normalidade dos dados do PC ROSA EVO. No entanto, permitiu concluir que recolhas dispersas ou concentradas após lavação podem alterar as decisões, tal como o controlo ser efetuado no momento antes de entrada na estufa ou após lavação. Na análise à saída da estufa os diferentes métodos não influenciam a medida de localização dos dados, mas observam-se diferentes variações entre valores individuais.

Pela análise do histórico de 2014, é expectável que em 2015 a distribuição das percentagens de humidade das estufas em 2015 não siga uma distribuição Normal. Em Janeiro de 2015, foram recolhidos 3070 valores individuais na Estufa 1 do tratamento ROSA EVO, 3010 valores individuais na Estufa 2, 2920 valores individuais na Estufa 3 e 3110 valores individuais na Estufa 4.

Pela análise das distribuições do **mês de Janeiro de 2015** nas estufas de tratamento ROSA EVO, histogramas da Figura 3.23 e Tabela 3.7, constata-se a obtenção de distribuições com elevados coeficientes de achatamento (superiores aos de 2014) e enviesamento positivo dos dados.

Estufa	Média	Desvio-padrão	Coef. de Assimetria	Coef. de Achatamento
d1	6.3245	1.4770	1.6425	8.7313
d2	6.2132	1.4496	1.5237	6.8553
d3	6.2394	1.3998	1.3052	5.5094
d4	6.6213	1.4898	1.1522	4.8061

Tabela 3.7: Média, Desvio-padrão, coeficiente de assimetria e Coeficiente de achatamento, dos registos de controlo em Janeiro de 2015

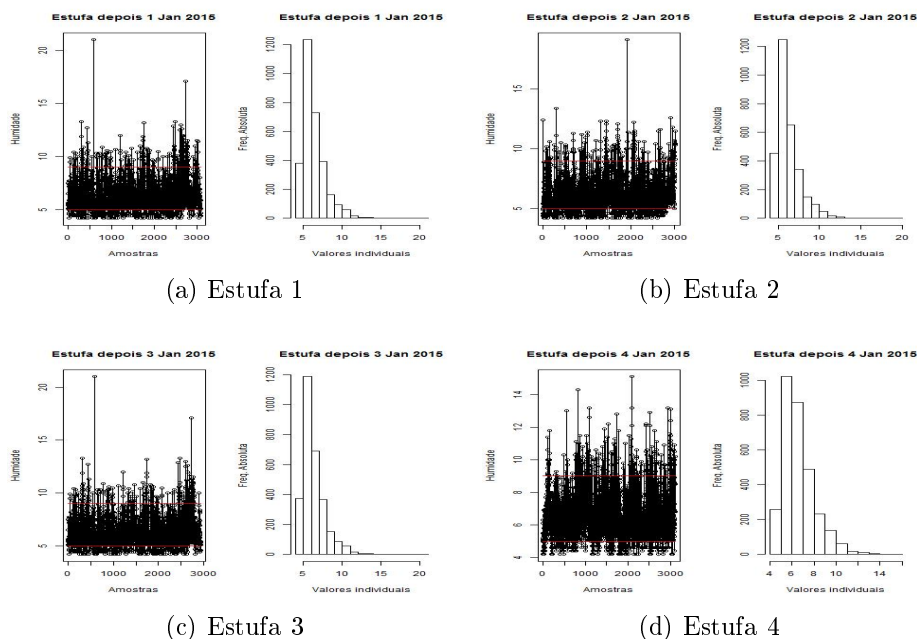


Figura 3.23: Histogramas dos dados de Janeiro 2015

Estufa/Estação	Mínimo	1º Quartil	2º Quartil	3º Quartil	Máximo	Amplitude
d1	4.2	5.4	6.0	7.0	21.0	16.8
d2	4.2	5.2	5.8	6.9	19.1	14.9
d3	4.2	5.3	5.9	6.9	15.2	11.0
d4	4.2	5.5	6.3	7.3	15.1	10.9

Tabela 3.8: Medidas de dispersão, dos registos de controlo em Janeiro de 2015

Através dos quartis dos valores individuais recolhidos nas diversas estufas, conclui-se que as percentagens de humidade são na sua maioria baixas. Ocorre uma grande amplitude de percentagens de humidade nas diversas estufas e a posição da mediana em comparação com a média reforça a conclusão do enviesamento positivo dos dados das estufas.

Na Tabela 3.8 constata-se que em todas as estufas obteve-se mínimo de 4.2% de humidade nos valores individuais recolhidos em Janeiro de 2015, que são um pouco dispares dos dados recolhidos em 2014. Esta limitação deve-se ao equipamento de medição de humidade. O novo equipamento utilizado somente recolhe valores superiores a 4%. Ao longo do ano de 2014 foram recolhidos apenas 96 valores individuais, o que se traduz em 0.06% dos dados recolhidos em 2014.

Baseada neste facto e pelos quartis presentes na Tabela 3.8 ignorou-se esta limitação do Aquaboy, para este estudo.

Tal como anteriormente, recorreu-se ao teste de Kolmogorov-Smirnov para testar a Normalidade dos dados de Janeiro 2015.

Estufa	Estatística de teste	$p\text{-value}$
d1	0.1207	$< 2.2 \times 10^{-16}$
d2	0.1283	$< 2.2 \times 10^{-16}$
d3	0.1051	$< 2.2 \times 10^{-16}$
d4	0.1014	$< 2.2 \times 10^{-16}$

Tabela 3.9: Resultado teste de Normalidade de Kolmogorov-Smirnov, Janeiro 2015

Com o resultado deste teste para um nível de significância de 5%, rejeita-se a hipótese nula, aceitando-se a não Normalidade das distribuições. Ou seja, como para todas as estufas obteve-se um $p\text{-value}$ menor que o nível de significância ($0.05 = \alpha > 2.2 \times 10^{-16} > p\text{-value}$) rejeitando-se a hipótese das distribuições seguirem uma distribuição Normal.

Contudo, assumindo os riscos inerentes ao ignorar a Normalidade dos dados, as representação das cartas $\bar{X} - R$ e $\bar{X} - S$, para Janeiro de 2015 são:

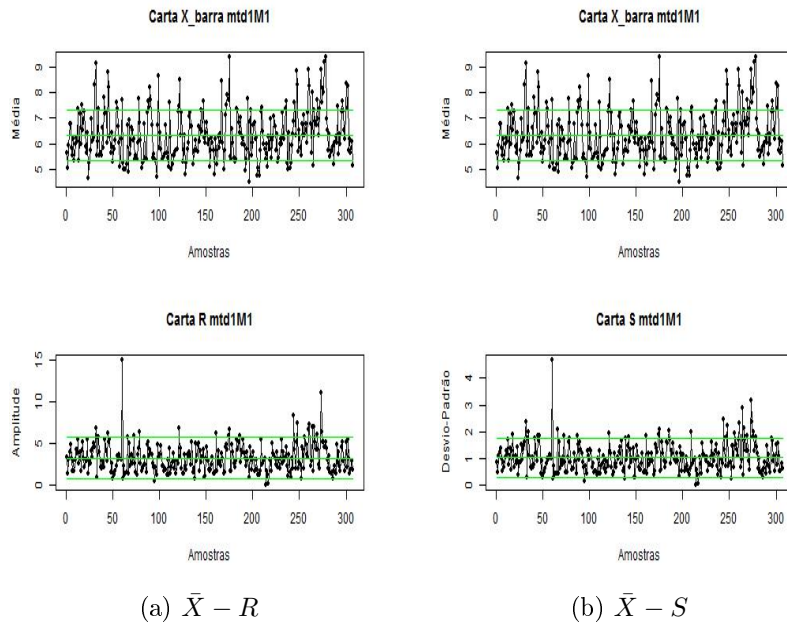


Figura 3.24: Carta de Controlo de Shewhart para estufa 1

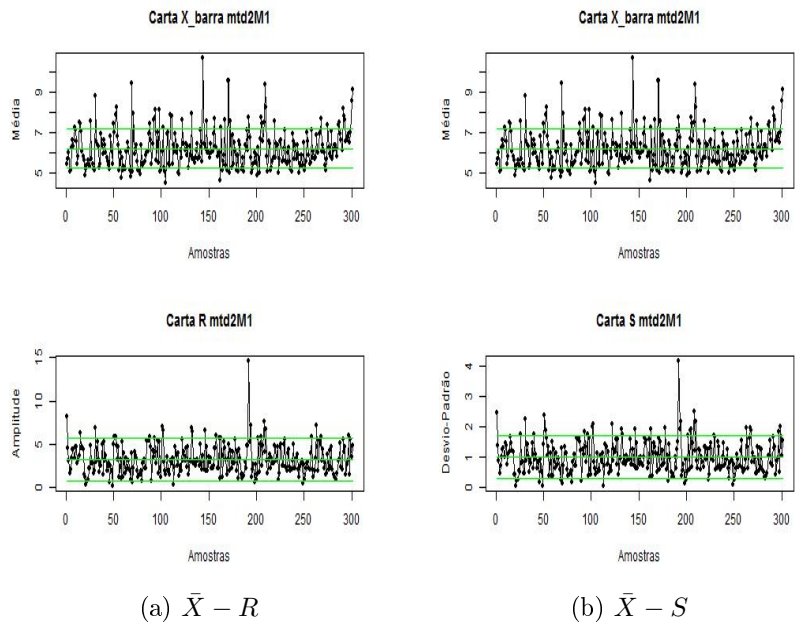


Figura 3.25: Carta de Controle de Shewhart para estufa 2

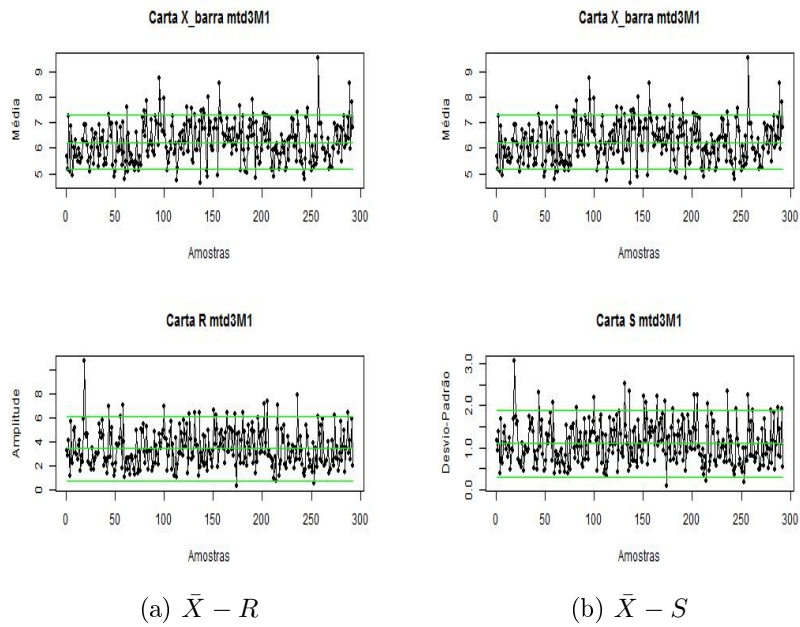


Figura 3.26: Carta de Controle de Shewhart para estufa 3

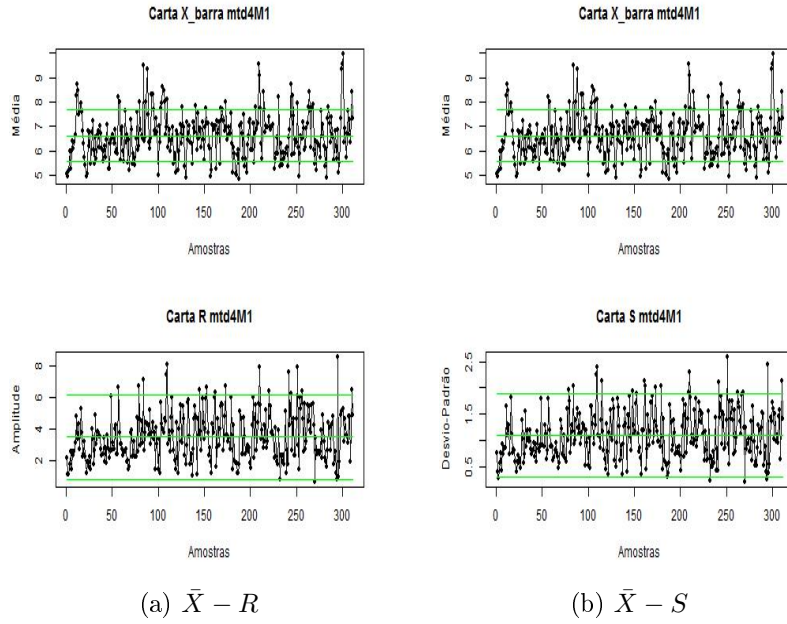


Figura 3.27: Carta de Controle de Shewhart para estufa 4

Limites de controle de Shewhart e número de não conformes, assumindo que as percentagens de humidade nas Estufas ROSA EVO seguem distribuição Normal, em Janeiro de 2015.

Estufa	Carta	Medida	LCS	LC	LCI	np LCS	np LCI
d1	$\bar{X} - R$	\bar{X}	7.3149	6.3245	5.3341	51	42
		R	5.7142	3.2156	0.7171	20	7
	$\bar{X} - S$	\bar{X}	7.3267	6.3245	5.3223	51	42
		S	1.7639	1.0279	0.2919	30	11
d2	$\bar{X} - R$	\bar{X}	7.1994	6.2132	5.2270	44	43
		R	5.6899	3.2020	0.7140	25	10
	$\bar{X} - S$	\bar{X}	7.1956	6.2132	5.2308	44	45
		S	1.7291	1.0076	0.2862	28	16
d3	$\bar{X} - R$	\bar{X}	7.2946	6.2394	5.1842	24	28
		R	6.0881	3.4260	0.7640	20	3
	$\bar{X} - S$	\bar{X}	7.3118	6.2394	5.1670	22	26
		S	1.887	1.0999	0.3128	21	3
d4	$\bar{X} - R$	\bar{X}	7.6929	6.6213	5.5496	44	44
		R	6.1829	3.4794	0.7759	17	1
	$\bar{X} - S$	\bar{X}	7.6960	6.6212	5.5465	44	44
		S	1.8916	1.1023	0.3131	18	4

Tabela 3.10: Limites de Controle e np , dados originais de Janeiro 2015

Alguns autores têm como referência para o estudo da Normalidade dos dados, não o comportamento dos valores individuais, mas sim o **comportamento das médias dos subgrupos**. Em seguida, serão analisadas as médias amostrais dos subgrupos:

Estufa/Estação	Mínimo	1º Quartil	2º Quartil	3º Quartil	Máximo	Amplitude
d1	4.49	5.62	6.14	6.93	9.41	4.92
d2	4.50	5.48	6.07	6.65	10.70	6.20
d3	4.63	5.66	6.16	6.82	9.57	14.94
d4	4.84	5.97	6.59	7.14	9.99	5.15

Tabela 3.11: Medidas de dispersão, das médias amostrais de Janeiro de 2015

Através das medidas anteriores conclui-se que as médias amostrais têm comportamento díspar dos valores individuais dos subgrupos. As médias amostrais têm uma menor variabilidade. Através dos quartis nota-se que a maioria dos valores médios estão localizados à esquerda do valor médio das médias amostrais.

Estufa	Média	Desvio-padrão	Coef. de Assimetria	Coef. de Achatamento
d1	6.3245	0.9787	0.7717	3.2620
d2	6.2132	0.9665	1.1847	5.0643
d3	6.2394	0.8044	0.5222	3.4530
d4	6.6213	0.9650	0.5712	3.4645

Tabela 3.12: Média, Desvio-padrão, coeficiente de assimetria e Coeficiente de achatamento, das médias amostrais de Janeiro de 2015

Os coeficientes de assimetria e achatamento salientam que os dados não correspondem a uma distribuição Normal.

Através dos gráficos seguintes, na estufa 2 é evidente o enviesamento positivo dos dados, as estufas 3 e 4 são as que não têm uma distribuição tão enviesada. Comparando a Figura 3.23 com 3.28, é notória a diferença do ajustamento à distribuição Normal dos valores individuais e das médias dos subgrupos.

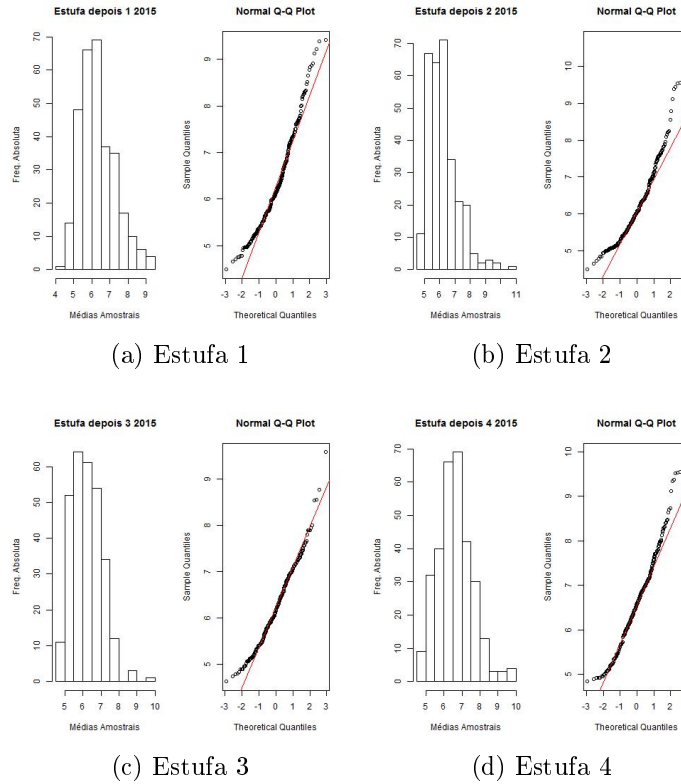


Figura 3.28: Distribuição das médias dos subgrupos de Janeiro 2015, das diversas Estufas ROSA EVO

Para validação da Normalidade das médias dos subgrupos do mês de Janeiro de 2015, realizou-se o teste de Kolmogorov-Smirnov. Para um nível de significância de 5%, através do teste de hipóteses de Kolmogorov-Smirnov não se rejeita a hipótese nula nas estufas 3 e 4. Tomando um nível de significância de 1%, rejeita-se a hipótese nula em todas as estufas, aceitando-se a Normalidade para todas as estufas. Portanto, para obter uma confirmação mais fiável realizou-se outro teste de Normalidade.

Estufa	Kolmogorov-Smirnov		Shapiro-Wilk	
	Estatística de teste	<i>p-value</i>	Estatística de teste	<i>p-value</i>
d1	0.0912	0.0122	0.9554	4.619×10^{-8}
d2	0.0936	0.0103	0.9257	4.259×10^{-11}
d3	0.0568	0.3032	0.9781	1.874×10^{-4}
d4	0.0568	0.2681	0.9754	3.550×10^{-5}

Tabela 3.13: Resultados do teste de Kolmogorv-Smirnov e Shapiro-Wilk, aplicado às médias dos subgrupos de Janeiro 2015

Segundo Razali and Wah (2011), o teste de hipóteses para avaliação da Normalidade mais eficaz em amostras de menor dimensão é o de Shapiro-Wilk, isto porque o teste de Shapiro-Wilk não necessita da estimação dos parâmetros (média e desvio-padrão) e também por não ser genérico como o teste de Kolmogorov-Smirnov, tornando-se mais poderoso. Assim, tendo como referência o teste de Shapiro-Wilk não é aceite a hipótese nula em todas as estufas de ROSA EVO, para um nível de significância de 1% e 5%.

É de realçar a importância da escolha das medidas estatísticas definidas para o controlo do processo produtivo. **A utilização da média dos subgrupos para monitorização do processo produtivo, é meramente aconselhada quando este não tem grande variabilidade.** Os valores individuais dos subgrupos podem não ser dispare entre si, no entanto as médias destes podem ser uniformes. Neste caso, a monitorização pelas médias traduz um processo produtivo com subgrupos homogêneos, que na realidade são valores individuais heterogêneos. Tendo em conta as constatações anteriores, as Estufas do ROSA EVO na AI encontram-se numa fase prematura, consequentemente o estudo avançará com análise dos valores individuais.

3.2.2.2 Transformação Box-Cox

Por forma, a contornar a limitação da não Normalidade dos dados foi usada a **transformação de Box-Cox**, nos valores individuais recolhidos em Janeiro de 2015, na tentativa da normalização dos dados.

Na estimação do parâmetro λ , foi usado o *package* ‘forecast’ do *software* R (Athanasopoulos et al., 2015). O parâmetro λ que melhor ajusta os dados das percentagens de humidade das várias estufas à distribuição normal são:

Estufa	d1	d2	d3	d4
λ	-1.00	-1.00	-1.00	-0.85

Tabela 3.14: Valores de λ para transformação dos dados

Aplicando a transformação Box-Cox aos dados, com os λ anteriores, as distribuições dos dados obtidas têm as seguintes características:

Estufa	Média	Desvio-padrão	Coef. de Assimetria	Coef. de Achatamento
d1	0.8346	0.0330	0.1182	2.7505
d2	0.8316	0.0332	0.2246	2.6788
d3	0.8326	0.0330	0.0565	2.6608
d4	0.9321	0.0422	0.0254	2.7366

Tabela 3.15: Média, Desvio-padrão, coeficiente de assimetria e coeficiente de achatamento após Transformação Box-Cox, Janeiro de 2015

Estufa/Estação	Mínimo	1º Quartil	2º Quartil	3º Quartil	Máximo	Amplitude
d1	0.7619	0.8148	0.8333	0.8571	0.9524	0.1905
d2	0.7619	0.8077	0.8276	0.8551	0.9476	0.2057
d3	0.7619	0.8113	0.8305	0.8551	0.9342	0.1723
d4	0.8291	0.9002	0.9304	0.9593	1.0590	0.2299

Tabela 3.16: Medidas de dispersão, dos registos de inspeção de Janeiro de 2015 transformados

A diminuição dos valores dos coeficientes de assimetria e achatamento da distribuição assumida pelos valores individuais originais é manifesta em relação aos dos dados transformados. Os dados obtidos com a Transformação de Box-Cox são mais similares a uma distribuição Normal que os dados originais. Como ilustrado nas seguintes representações gráficas:

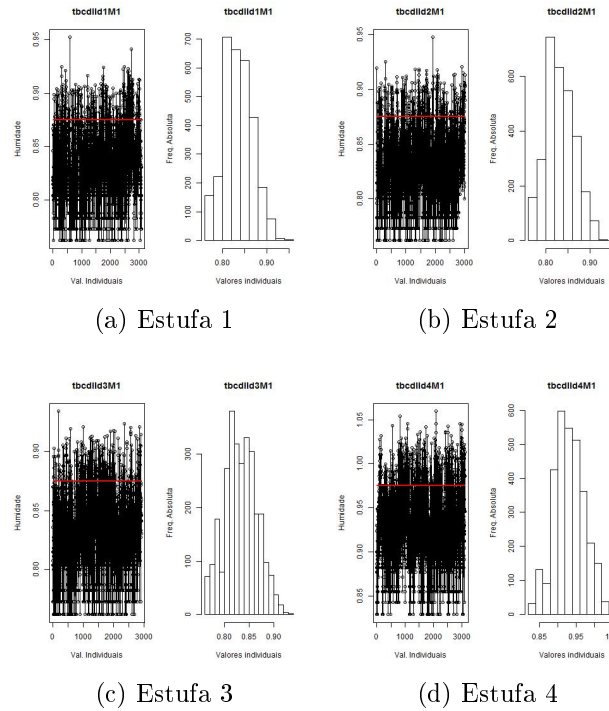


Figura 3.29: Valores individuais dos dados transformados de Janeiro 2015

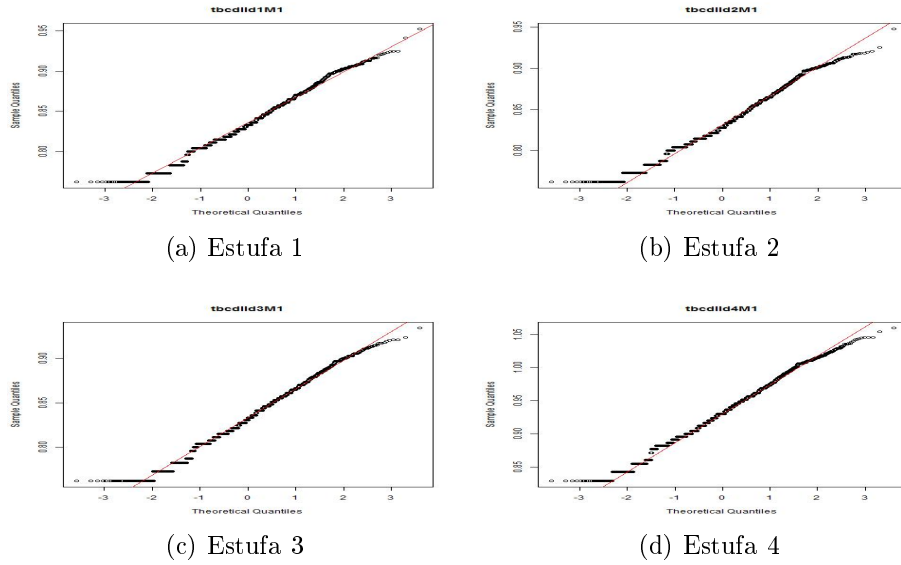


Figura 3.30: Ajustamento à distribuição Normal dos dados transformados de Janeiro 2015

Realizado o teste para a Normalidade de Kolmogorov-Smirnov aos dados de Janeiro transformados obteve-se:

Estufa	Estatística de teste	<i>p-value</i>
d1	0.0528	7.215×10^{-8}
d2	0.0735	1.532×10^{-14}
d3	0.0497	1.105×10^{-16}
d4	0.0385	1.951×10^{-4}

Tabela 3.17: Resultados teste de Kolmogorov-Smirnov, dados transformados Janeiro 2015

A hipótese nula não é aceite para as percentagens de humidade nas diversas estufas após a transformação dos dados.

Após a determinação do λ para a transformação de Box-Cox é necessário determinar os Limites de Especificação tendo em conta este parâmetro. Pela equação 2.35 os novos Limites de Especificação para as estufas 1, 2 e 3 em Janeiro 2015 são: $LIE = \frac{4^{-1}-1}{-1} = 0.750$ e $LSE = \frac{8^{-1}-1}{-1} = 0.875$. Para a estufa 4 tem-se os seguintes limites de especificação $LIE = \frac{4^{-0.85}-1}{-0.85} = 0.814$ e $LSE = \frac{8^{-0.85}-1}{-0.85} = 0.976$.

Pelos dados da Tabela 3.15 assume-se que os dados transformados seguem uma distribuição aproximadamente Normal. Deste modo, calculando as Cartas de Controle de Shewhart para os dados transformados por Box-Cox obtém-se:

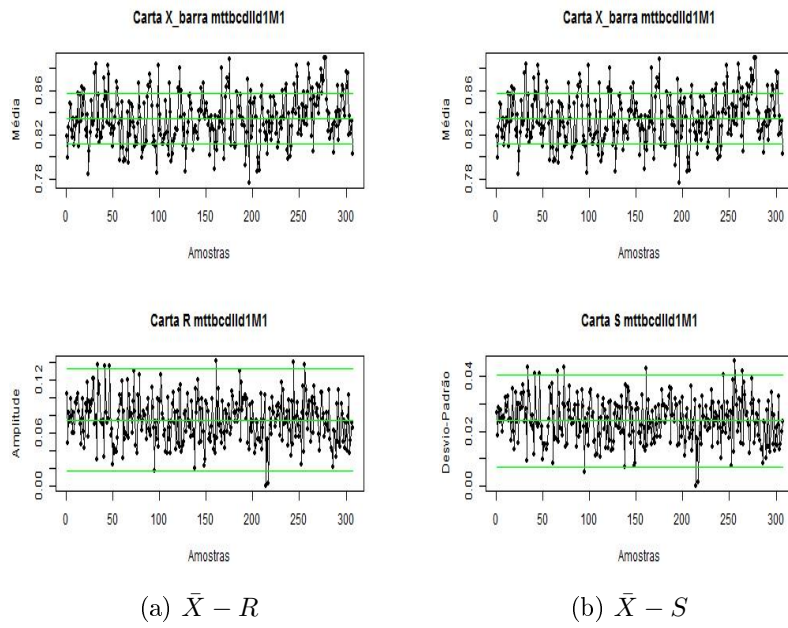


Figura 3.31: Carta de Controle de Shewhart para estufa 1

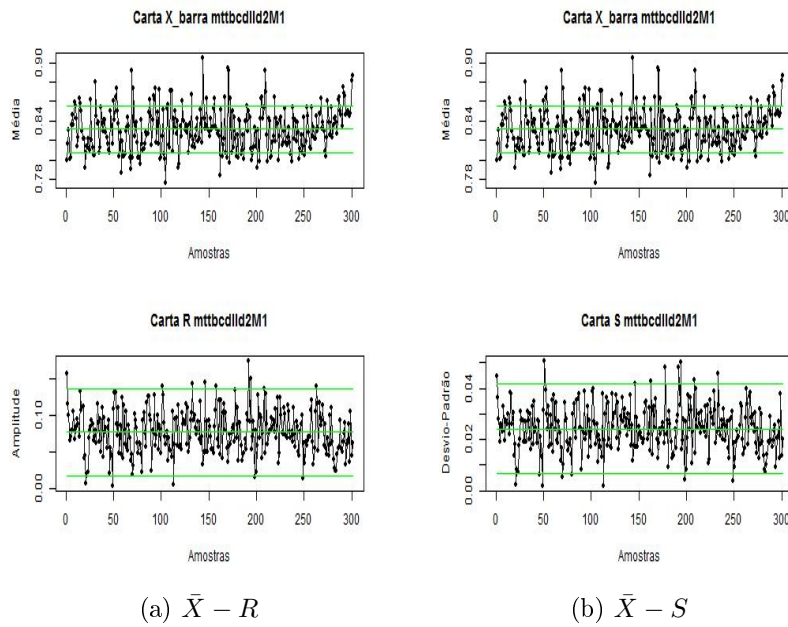


Figura 3.32: Carta de Controle de Shewhart para estufa 2

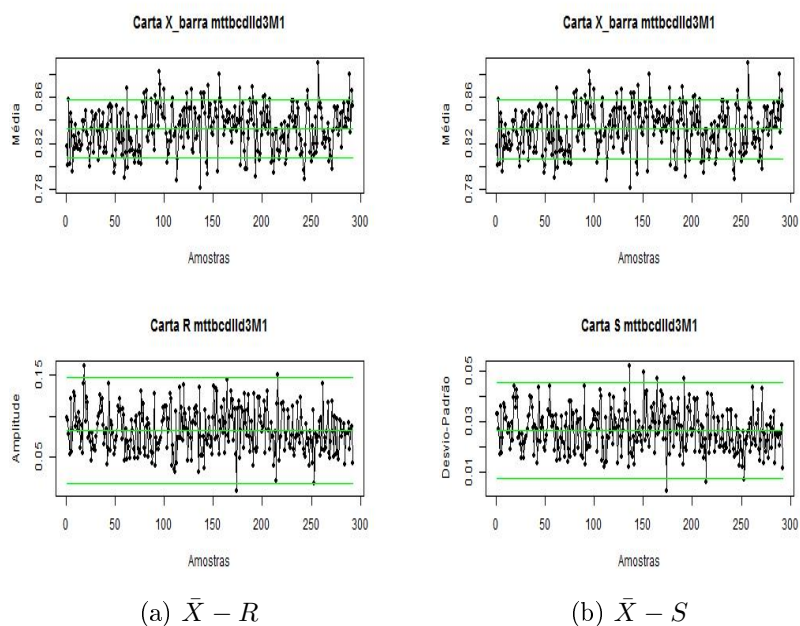


Figura 3.33: Carta de Controle de Shewhart para estufa 3

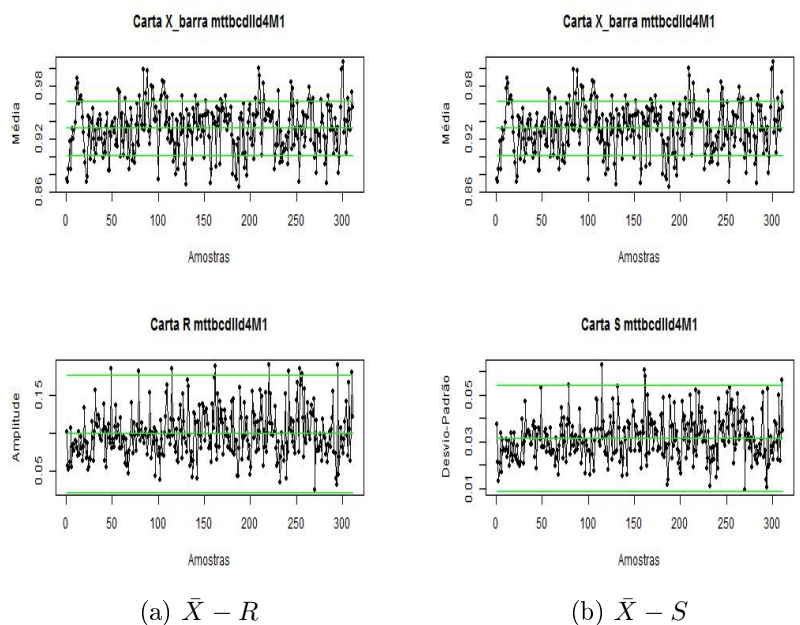


Figura 3.34: Carta de Controle de Shewhart para estufa 4

Limites de controle de Shewhart e número de não conformes, nas Estufas ROSA EVO para as percentagens de humidade em Janeiro de 2015 após a transformação

de Box-Cox.

Estufa	Carta	Medida	LCS	LC	LCI	np LCS	np LCI
d1	$\bar{X} - R$	\bar{X}	0.8575	0.8346	0.8116	57	52
		R	0.1323	0.0744	0.0166	6	2
	$\bar{X} - S$	\bar{X}	0.8578	0.8346	0.8114	56	52
		S	0.0408	0.0238	0.0068	8	3
d2	$\bar{X} - R$	\bar{X}	0.8553	0.8316	0.8079	42	50
		R	0.1368	0.0770	0.0172	8	5
	$\bar{X} - S$	\bar{X}	0.8553	0.8316	0.8080	42	50
		S	0.0417	0.0243	0.0069	9	8
d3	$\bar{X} - R$	\bar{X}	0.8580	0.8326	0.8073	25	35
		R	0.1462	0.0823	0.0184	2	2
	$\bar{X} - S$	\bar{X}	0.8584	0.8326	0.8069	24	35
		S	0.0453	0.0264	0.0075	4	3
d4	$\bar{X} - R$	\bar{X}	0.9627	0.9321	0.9014	44	50
		R	0.1769	0.0995	0.0222	10	0
	$\bar{X} - S$	\bar{X}	0.9628	0.9321	0.9013	44	50
		S	0.0541	0.0316	0.0090	5	0

Tabela 3.18: Limites de Controlo e número amostragens fora de controlo, Janeiro 2015

Comparando as cartas de controlo 3.24, 3.25, 3.26 e 3.27 com 3.31, 3.32, 3.33 e 3.34, respetivamente é evidente a elevada disparidade dos valores de np . Avaliando as cartas para a análise da localização do processo (carta \bar{X}) ocorre um ligeiro aumento do número de subgrupos fora de controlo estatístico. Não obstante, quando avaliada a variabilidade do processo produtivo (carta R e S) existe uma enorme redução do np .

3.2.2.3 Índices de Capacidade

Aqui serão avaliados os índices de capacidade do processo, para distribuições Normais e Não Normais, para os dados de humidade do mês de Janeiro de 2015, nas diversas estufas do ROSA EVO. Atualmente, na empresa o *SPC* utiliza exclusivamente os índices de capacidade para distribuições Normais, o que pela análise anterior é incorreto.

No caso dos índices de capacidade para distribuições Normais, a variabilidade inerente ao processo pode ser estimada, como descrito pelas equações 2.16 e 2.17. Dado que a dimensão do subgrupo é 10, pelas tabelas de constantes para Cartas de Controlo de Shewhart têm-se: $d_2 = 3.0780$ e $c_4 = 0.9727$.

Estufa	$\hat{\sigma}_0 = \frac{\bar{R}}{d_2}$	$\hat{\sigma}_0 = \frac{\bar{S}}{c_4}$	Real
d1	1.0447	1.0567	1.4770
d2	1.0403	1.0359	1.4496
d3	1.1131	1.1308	1.3998
d4	1.1304	1.1332	1.4898

Tabela 3.19: Desvio-padrão associado às estufas ROSA EVO, Janeiro 2015

Como é visível pela tabela anterior, na realidade a dispersão dos valores individuais de humidade é maior na estufa 4 e menor na estufa 3. No entanto, pela estimação dos desvios-padrões as estufas com maior variabilidade são a 3 e a 4.

Assim, assumindo que os dados originais são Normais, os índices de capacidade do processo são:

Estufa	Desvio-Padrão	C_p	C_{pk}	C_{pm}	C_{pmk}
d1	\bar{R}	0.6381	0.5346	0.6094	0.5106
	\bar{S}	0.6309	0.5285	0.6031	0.5052
	Real	0.4514	0.3781	0.4408	0.3693
d2	\bar{R}	0.6409	0.5725	0.6278	0.5609
	\bar{S}	0.6436	0.5750	0.6304	0.5631
	Real	0.4599	0.4109	0.4550	0.4065
d3	\bar{R}	0.5989	0.5273	0.5856	0.5155
	\bar{S}	0.5896	0.5190	0.5768	0.5077
	Real	0.4763	0.4193	0.4695	0.4133
d4	\bar{R}	0.5898	0.4066	0.5168	0.3563
	\bar{S}	0.5883	0.4055	0.5159	0.3556
	Real	0.4475	0.3085	0.41307	0.2847

Tabela 3.20: Índices de Capacidade do Processo para distribuições Normais, dados originais Janeiro 2015

Os índices de capacidade calculados pelo desvio-padrão dos valores individuais recolhidos no mês de Janeiro são inferiores aos índices de capacidade através da estimação do desvio-padrão pelos fatores tabelados das cartas de controlo de Shewhart. Pela análise da tabela anterior, observa-se que nenhuma das estufas é capaz de cumprir as especificações de humidade. Apesar disto, é a estufa 4 que obtém o pior desempenho. Em todas as situações, os índices C_p e C_{pk} são bastantes baixos, pelo que se conclui que o processo produtivo tem uma maior variabilidade que a permitida e não existe uma localização do processo produtivo no valor nominal.

Como constatado, na secção 2.2.3 não podem ser usados estes índices de capacidade devido à Não Normalidade das distribuições das estufas ROSA EVO. Portanto,

seguidamente foram estimados os índices de capacidade para distribuições Não Normais. Serão calculados por dois dos métodos apresentados. Para tal têm de ser calculados os percentis associados a cada distribuição dos valores individuais da percentagem de humidade das diversas estufas.

Estufa	$F_{0.135}$	F_{50}	$F_{99.865}$
d1	4.2000	6.0000	13.1714
d2	4.2000	5.8000	12.3000
d3	4.2000	5.9000	12.6060
d4	4.2000	6.3000	13.1803

Tabela 3.21: Percentis associados às estufas ROSA EVO, Janeiro 2015

Em todas as estufas o percentil 0.135% é igual. Sendo que os percentis 50% e 99.865% variam consoante a estufa, o que é expectável devido ao coeficiente de assimetria apresentados em 3.7.

Estufa	Método	C_p	C_{pk}	C_{pm}	C_{pmk}
d1	Clements	0.4459	0.2789	0.4459	0.2789
	Pearn and Chen	0.4459	0.4459	0.4459	0.4459
d2	Clements	0.4938	0.3385	0.4885	0.3370
	Pearn and Chen	0.4938	0.4444	0.4885	0.4396
d3	Clements	0.4759	0.3132	0.4746	0.3128
	Pearn and Chen	0.4759	0.4521	0.4746	0.4509
d4	Clements	0.4454	0.2471	0.4367	0.2450
	Pearn and Chen	0.4454	0.3786	0.4367	0.3712

Tabela 3.22: Índices de Capacidade do Processo para distribuições Não Normais, Janeiro 2015

A capacidade do processo produtivo quando avaliada pelos índices de capacidade apropriados demonstram diferenças em relação aos índices para distribuições Normais. Pelos atuais índices de capacidade conclui-se que o processo produtivo é incapaz. Pelos índices adequados obteve-se um processo menos capaz.

Os índices de capacidade que têm em conta apenas a medida de dispersão indicam um processo produtivo mau, mas pelos índices que levam em atenção a variabilidade e localização do processo o estado deste é ainda pior que o concluído anteriormente.

Para os dados transformados por Box-Cox também foram calculados os índices de capacidade para dados Normais:

Estufa	$\hat{\sigma}_0 = \frac{\bar{R}}{d_2}$	$\hat{\sigma}_0 = \frac{\bar{S}}{c_4}$	Real
d1	0.0242	0.0245	0.0330
d2	0.0250	0.0250	0.0332
d3	0.0267	0.0271	0.0330
d4	0.0323	0.0324	0.0422

Tabela 3.23: Desvio-padrão dados transformados, Janeiro 2015

Estufa	Desvio-Padrão	C_p	C_{pk}	C_{pm}	C_{pmk}
d1	\bar{R}	0.8614	0.5573	0.6363	0.4117
	\bar{S}	0.8515	0.5509	0.6323	0.4090
	Real	0.6320	0.4089	0.5252	0.3398
d2	\bar{R}	0.8329	0.5779	0.6615	0.4590
	\bar{S}	0.8347	0.5792	0.6624	0.4596
	Real	0.6270	0.4350	0.5433	0.3770
d3	\bar{R}	0.7792	0.5281	0.6224	0.4219
	\bar{S}	0.7676	0.5203	0.6165	0.4179
	Real	0.6313	0.4279	0.5389	0.3653
d4	\bar{R}	0.8349	0.4527	0.5487	0.2976
	\bar{S}	0.8323	0.4513	0.5480	0.2972
	Real	0.6396	0.3468	0.4805	0.2606

Tabela 3.24: Índices de Capacidade do Processo para distribuições Normais, dados transformados Janeiro 2015

Através dos índices de capacidade para os dados transformados, conclui-se que o processo nas Estufas de ROSA EVO não é capaz, dado os seus baixos índices C_p nas várias estufas. Os índices de capacidade C_{pk} não são bons (como plausível), indicando que a humidade das rolhas após o ROSA EVO não é capaz de estar centrada nas especificações impostas ao processo produtivo.

3.2.2.4 Análise de Correlação

O processo produtivo no ROSA EVO deverá ser adaptado para o cumprimento das especificações impostas. Para averiguar qual a relação dos valores de humidade antes e após o ROSA EVO, foi efetuada uma análise²² destas amostragens, levando em consideração as principais variáveis de programação deste PC.

Inicialmente os registos em *SPC* continham somente as informações: OF, tipo de rolha, calibre, classe e lavação. Com estes, não era possível uma identificação

²²A análise foi efetuada com recurso ao *software IBM SPSS Statistics*, versão 21.

inequívoca dos lotes em histórico. Consequentemente, foram adicionados ao *software* os campos: N (identificativo da amostra coletada para quantificação de TCA), quantidade da OF, destino e observações. Assim, a atual configuração dos registos permite um rastreio sem ruído.

Nesta análise²³ apenas foram considerados os registos que continham todos estes campos preenchidos. Por conseguinte não foram consideradas bastantes amostragens do ROSA EVO 2. Foram avaliadas 2980 “transações”, sendo que 863 passaram no ROSA EVO 1, 360 no ROSA EVO 2, 848 no ROSA EVO 3 e 909 no ROSA EVO 4.

Nos seguintes histogramas estão representadas todas as médias amostrais obtidas em função do ROSA EVO por qual os lotes foram passados.

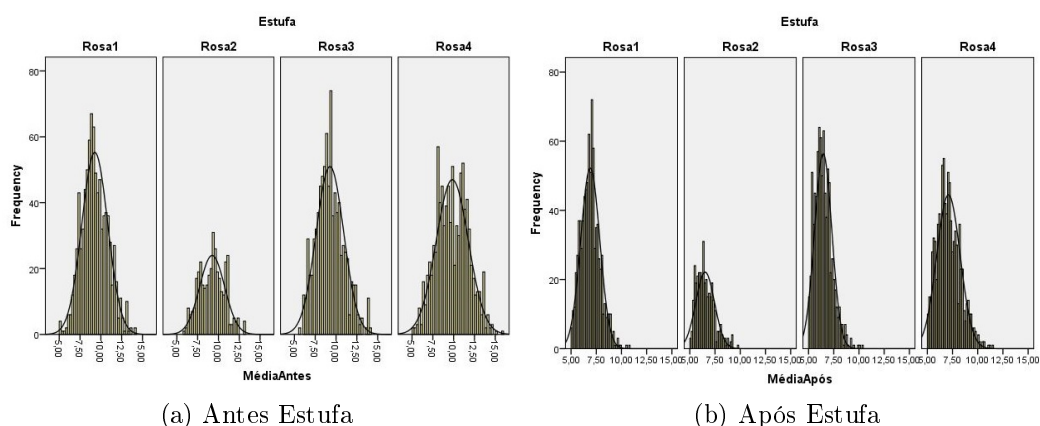


Figura 3.35: Distribuição dos valores médios das amostragens

Para avaliar a hipótese das variáveis seguirem uma distribuição Normal, foram efetuados os testes estatísticos de ajustamento Kolmogorov-Smirnov e Shapiro-Wilk, com recurso ao *software SPSS*²⁴.

Por ambos os testes conclui-se que as médias amostrais não seguem uma distribuição Normal, para um nível de significância de 5%. Ou seja, em todas as estufas o *p-value* é inferior a 0.05, tanto antes como após estufa.

²³Para avaliar uma correspondência entre as humidades percentuais antes e após estufa, caso exista mais que uma amostragem após estufa, considera-se que as amostragens antes estufa são todas iguais à recolhida antes estufa.

²⁴Visto que o teste de Kolmogorov-Smirnov necessita da estimação dos parâmetros da distribuição é referida a correção de significância de Lilliefors.

Tests of Normality							
Estufa	Kolmogorov-Sm nov ^a			Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
MédiaAntes	Rosa1	,043	863	,001	,994	863	,001
	Rosa2	,053	360	,018	,989	360	,007
	Rosa3	,055	848	,000	,984	848	,000
	Rosa4	,045	909	,000	,993	909	,000

a. Lilliefors Significance Correction

Tests of Normality							
Estufa	Kolmogorov-Sm irnov ^a			Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
MédiaApós	Rosa1	,052	863	,000	,987	863	,000
	Rosa2	,074	360	,000	,964	360	,000
	Rosa3	,056	848	,000	,963	848	,000
	Rosa4	,054	909	,000	,979	909	,000

a. Lilliefors Significance Correction

(a) Antes Estufa

Tests of Normality							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
MédiaApós	Rosa1	,052	863	,000	,987	863	,000
	Rosa2	,074	360	,000	,964	360	,000
	Rosa3	,056	848	,000	,963	848	,000
	Rosa4	,054	909	,000	,979	909	,000

a. Lilliefors Significance Correction

(b) Após Estufa

Figura 3.36: Teste Normalidade para média amostrais de humidade

Ocorreu uma redução das percentagens de humidade nos lotes após a passagem pelos ROSA EVO. Para uma melhor compreensão desta redução, foram elaborados os gráficos de dispersão dos valores médios e desvios padrão das amostragens recolhidas antes e após estufa, em função desta. O grau de associação entre variáveis pode ser medido com recurso ao **coeficiente de correlação**. Os mais utilizados são os coeficientes de **Pearson** e **Spearman**. O coeficiente de correlação linear de Pearson apenas é válido para variáveis que seguem uma distribuição Normal. Uma vez que, ambas as medidas em estudo não seguem uma distribuição Normal, os coeficientes de correlação foram calculados através do coeficiente de Spearman.

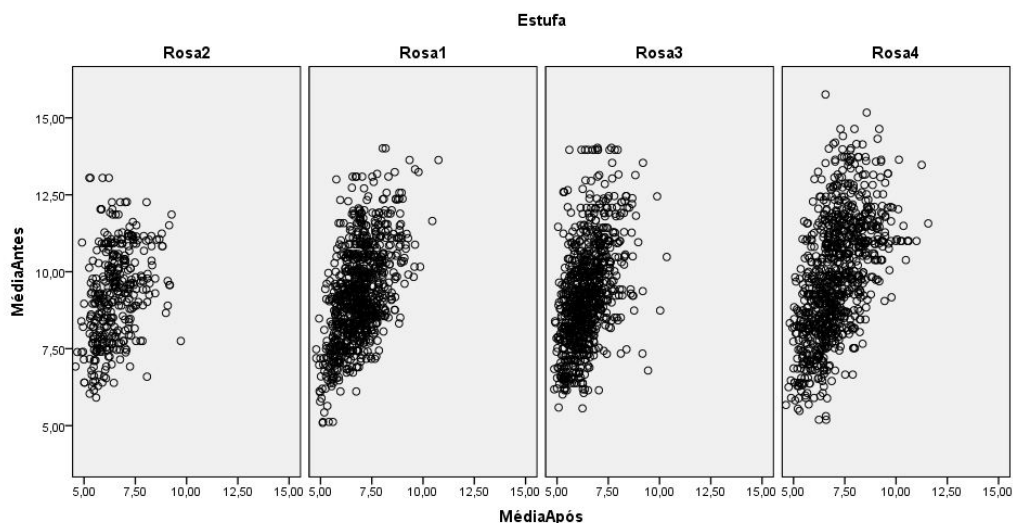


Figura 3.37: Gráfico de dispersão da média amostral antes *versus* após ROSA EVO

É aparente uma relação positiva entre as humidades médias das amostras antes e após estufa. Isto é, se a humidade percentual do lote após lavagem for elevada, a humidade percentual após o ROSA EVO também será elevada, apesar de ser menor que a inicial.

	Estufa 1		Estufa 2		Estufa 3		Estufa 4	
	Antes	Depois	Antes	Depois	Antes	Depois	Antes	Depois
Antes	1.000	0.504	1.000	0.406	1.000	0.502	1.000	0.582
Depois	0.504	1.000	0.406	1.000	0.502	1.000	0.582	1.000

Tabela 3.25: Coeficiente Spearman entre Média amostral antes e após ROSA EVO

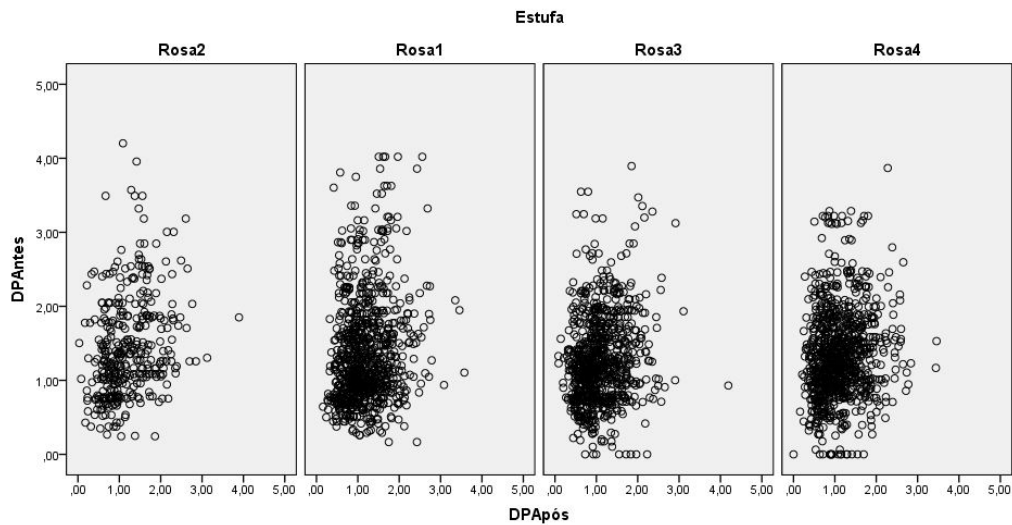


Figura 3.38: Gráfico de dispersão do desvio-padrão amostral antes *versus* após ROSA EVO

Na Figura 3.38 o “efeito nuvem” indica a inexistência de forte correlação entre as variáveis.

	Estufa 1		Estufa 2		Estufa 3		Estufa 4	
	Antes	Depois	Antes	Depois	Antes	Depois	Antes	Depois
Antes	1.000	0.167	1.000	0.324	1.000	0.191	1.000	0.156
Depois	0.167	1.000	0.324	1.000	0.191	1.000	0.156	1.000

Tabela 3.26: Coeficiente Spearman entre desvio-padrão amostral antes e após ROSA EVO

Tal como esperado pelas imagens 3.37 e 3.38, o coeficiente de correlação é superior entre as médias amostrais antes e após estufas ROSA EVO.

Nas várias estufas foram estudadas as variáveis principais na programação de cada estufa. As variáveis que não estão inertes no espaço temporal em estudo foram:

- **Estufa ROSA EVO 1:** Variável B assumiu 2 valores diferentes.

Através do coeficiente de correlação de Spearman conclui-se que não existe interligação da variável B com os resultados das amostragens, pois todos os coeficientes são inferiores ou iguais a 0.254 (fraca correlação).

- **Estufa ROSA EVO 2:** Variável B assumiu 2 valores diferentes.

O cenário na estufa 2 em termos de interligações de variáveis é muito fraco (coeficientes de correlação de Spearman são baixos inferiores a 0.100).

- **Estufa ROSA EVO 3:** Variável A, B, C, D e F assumem 2 valores, no entanto não necessariamente nas mesmas amostragens.

O coeficiente de correlação entre as variáveis A, C, D e F é de 1.00. Ou seja, basta analisar uma das variáveis que as restantes terão um comportamento semelhante a estas. Também para esta estufa os coeficientes de correlações estão abaixo dos 0.500, o que traduz uma relação fraca entre variáveis, onde não é possível aferir conclusões sobre estas.

- **Estufa ROSA EVO 4:** Variável B, C, D e F assumem 2 valores, no entanto não necessariamente nas mesmas amostragens.

Na estufa 4 as variáveis B, C e E variaram em concordância, desta forma a coeficiente de correlação assume a valor 1 (correlação perfeita). Os restantes coeficientes de correlação indicam uma fraca correlação entre variáveis.

Assim, nada se pode concluir acerca da influência das variáveis parametrizáveis do ROSA EVO na redução da humidade percentual dos lotes. Os coeficientes de correlação mais elevados são entre a média amostral antes e após estufa, apresentados na Tabela 3.25.

3.2.3 Cartas de Controlo de EWMA

As **cartas de controlo EWMA** foram estudadas para este PC, dado que estas não são vulneráveis à hipótese de Normalidade dos dados e pela eficiência na deteção de pequenas alterações. Tal como na análise das cartas de controlo de Shewhart o foco de análise serão as percentagens de humidade em **Janeiro de 2015**, nas várias estufas ROSA EVO. Para a construção destas cartas de controlo foi usado o *package* ‘qcc’²⁵, disponível no *Software R* (Scrucca, 2015).

Na secção 2.2 foi mencionado que as cartas de controlo de EWMA dependem de duas constantes: constante de amortecimento (λ) e fator de multiplicação dos desvios-padrão nos limites de controlo (L). Segundo Montgomery (1996) estes valores devem situar-se nos seguintes intervalos: $0.05 \leq \lambda \leq 0.25$ e $2.6 \leq L \leq 2.8$.

²⁵Nas cartas de controlo EWMA, os pontos • representam as estatísticas de EWMA, calculadas com recurso à equação 2.30, e os pontos + correspondem às médias dos subgrupos.

Numa primeira fase construíram-se as cartas de controlo EWMA com o valor de $L = 2.998$ (conceito carta $3 - \sigma$) e $\lambda = 0.25$, para todas as estufas ROSA EVO, obtendo-se:

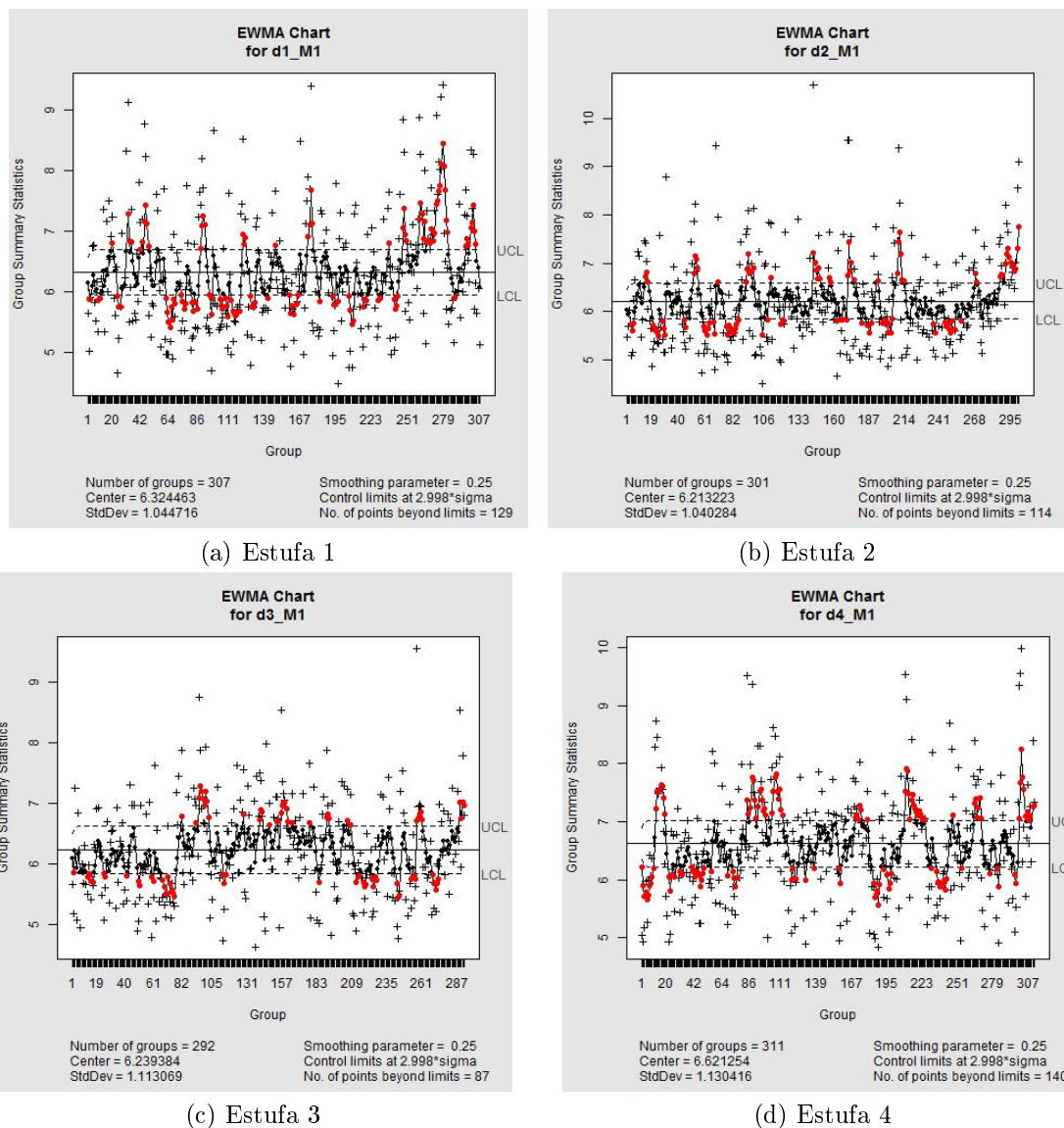


Figura 3.39: Cartas de Controlo EWMA para $\lambda = 0.25$ e $L = 2.998$

A LC da carta de controlo EWMA é igual à da carta \bar{X} de Shewhart, já que não depende de λ nem de L . Pelas cartas EWMA infere-se que o processo produtivo não está centrado no valor nominal da especificação, o que corrobora as conclusões retiradas pelos índices de capacidade (processo produtivo não capaz neste PC).

A amplitude dos limites de controle de EWMA é mais reduzida que na carta \bar{X} de Shewhart. Porém, trata-se de uma carta de controle de médias móveis exponencialmente ponderadas sendo necessária atenção na sua análise.

	Shewhart		EWMA	
	Número np	Primeiro np	Número np	Primeiro np
Estufa 1	93	13	129	2
Estufa 2	87	9	114	4
Estufa 3	52	44	87	2
Estufa 4	88	11	140	1

Tabela 3.27: Não Conformidades carta de controle \bar{X} e EWMA

Na carta de controle EWMA é evidente um maior número de amostragens fora dos limites de controle. A maioria das amostragens que se encontravam fora dos limites de controle nas cartas de Shewhart, também estavam nas cartas EWMA. As amostragens fora de controle são disputadas antecipadamente, o que torna possível uma atuação prévia que poderá prevenir que o processo produtivo esteja fora de controle estatístico.

Analisando as cartas EWMA para os valores referência apresentados em Lucas and Saccucci (1990) foram obtidos os seguintes resultados:

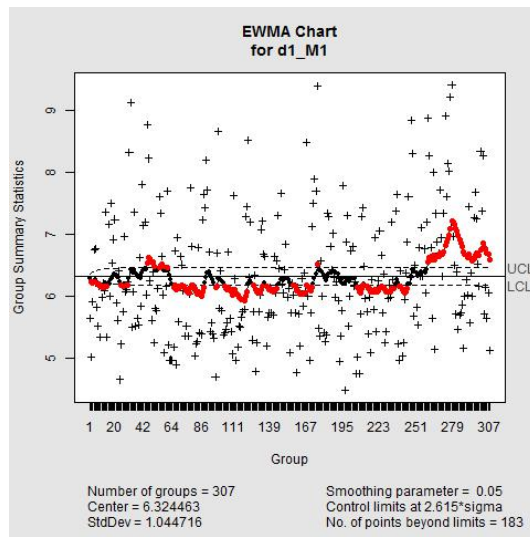
	$\lambda = 0.40$ $L = 3.054$		$\lambda = 0.20$ $L = 2.962$		$\lambda = 0.10$ $L = 2.814$		$\lambda = 0.05$ $L = 2.615$	
	Total	Primeiro	Total	Primeiro	Total	Primeiro	Total	Primeiro
Estufa 1	116	2	138	2	157	2	183	2
Estufa 2	111	4	112	4	116	4	112	2
Estufa 3	77	2	94	2	117	2	163	2
Estufa 4	116	1	145	1	165	1	168	7

Tabela 3.28: Não Conformidades para valores de referência de λ e L , na carta de controle EWMA de Janeiro de 2015

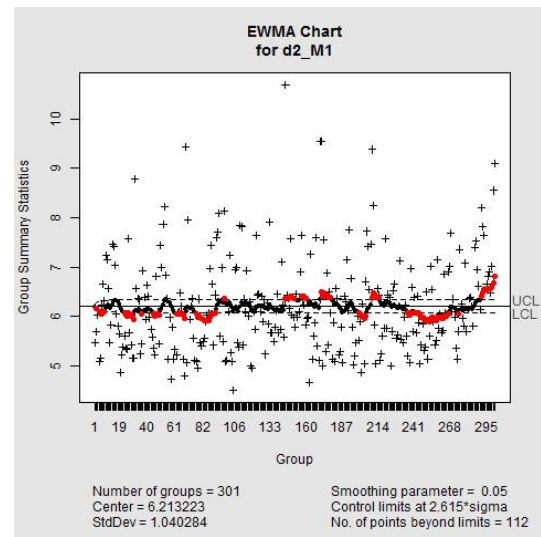
Para todos os valores de referência usados, os alarmes são gerados anteriormente comparativamente às cartas de controle de Shewhart.

Como fica evidente pela Tabela 3.28, há medida que o valor de L e λ diminui o número de np aumenta, isto deve-se:

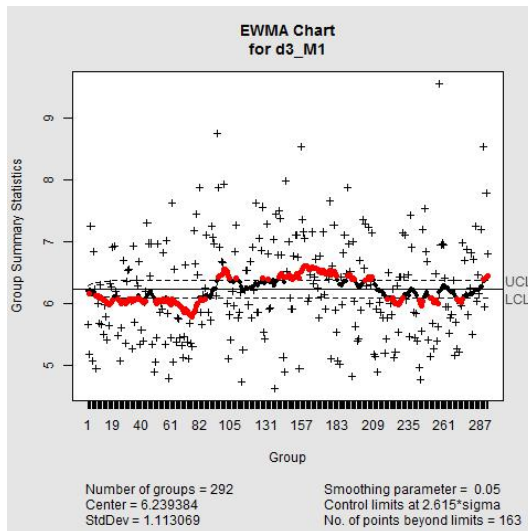
- À diminuição da distância entre o LCI e o LCS;
- Ao subgrupo mais recente perder peso, pelo que são os subgrupos mais antigos que começam a pesar mais na estatística de EWMA atual.



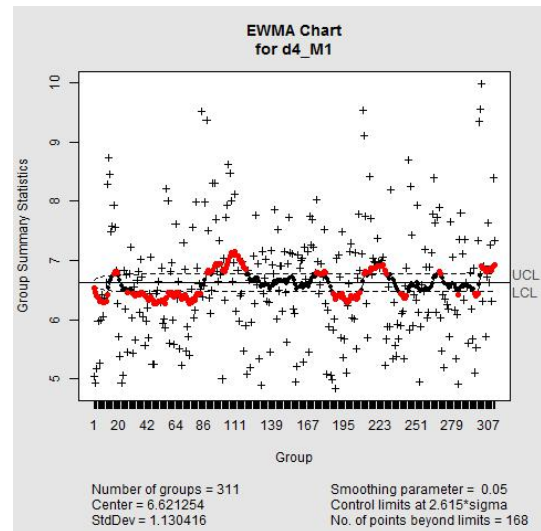
(a) Estufa 1



(b) Estufa 2



(c) Estufa 3



(d) Estufa 4

Figura 3.40: Cartas de Controle EWMA para $\lambda = 0.05$ e $L = 2.615$

As cartas EWMA dão uma melhor noção do processo produtivo, dado que detetam as pequenas variações que podem levar a um processo produtivo incapaz. Com estas também conclui-se que o processo produtivo nas várias estufas ROSA EVO não está centrado no valor nominal imposto.

Capítulo 4

Conclusões e Trabalho Futuro

4.1 Conclusões

O OLTP é um sistema rudimentar em que é necessário um conhecimento pormenorizado das tabelas de armazenamento de dados. Este tipo de SI é bom para extrair informação de históricos reduzidos, contudo é propício a consultas com erros. Enquanto que o OLAP é um SI estruturado, coeso, dinâmico e fiável, que pressupõe um grande trabalho de *back-End* para que a informação neste seja despoletada sobre várias perspetivas. Por conseguinte, a estruturação do OLAP é crucial para o sucesso deste.

Neste trabalho foi possível corrigir erros, introduzir adaptações/melhorias no BI com os dados da Qualidade da Amorim & Irmãos S.A., tal como nos seus *softwares* de fonte de dados. Isto permitiu que os utilizadores obtivessem a informação de uma forma mais segura e completa, que abastece as análises fundamentais para o conhecimento dos produtos e processos produtivos, o que anteriormente não era possível.

O QUALIS para além de auxiliar o departamento da Qualidade, também permite a partilha de informação com os restantes departamentos. Dada a estrutura do cubo ele é intuitivo, assim os utilizadores não necessitam de conhecer profundamente o funcionamento dos *softwares*, como é exigido num SI OLTP, para efetuar consultas da informação. Assim, os restantes departamentos também beneficiam do QUALIS, pois têm a informação da Qualidade atualizada para dar seguimentos às tarefas.

O OLTP não é pior que o OLAP, uma vez que são vocacionados para objetivos distintos. Ambos têm vantagens e desvantagens. Sumariamente, cabe a cada empresa optar pelo SI que mais lhe é útil ou até mesmo aliar os dois.

No âmbito do CEP, a empresa ainda se encontra num estágio inicial. No caso de estudo real, foi detetado que nas estufas ROSA EVO já existiam os limites de controlo pré-estabelecidos para as cartas de \bar{X} : $LSC = 10,0\%$ e $LIC = 6,0\%$.

Estes têm de ser necessariamente retificados, dado que ao assumir $LCS > LES$ existe um grande risco, porque à partida já é aceite uma percentagem considerável de produtos não conformes. O ideal é que o processo produtivo tenha a amplitude dos limites de controlo inferior à dos limites de especificação, em que a LC coincida com o valor nominal. No PC ROSA EVO, as amplitudes são iguais e para além disso os limites de controlo pré-estabelecidos não variam em função da estação do ano. Deste modo, nas estações Primavera e Verão a percentagem de aceitação de produtos não conformes é inferior às estações Outono e Inverno. Sendo que, no Outono e Inverno o valor nominal dos limites de especificação é coincidente com o LCI.

Na análise da Normalidade da humidade, conclui-se que no ROSA EVO os dados não seguem uma distribuição Normal. Assim, as cartas de controlo de Shewhart e índices de capacidade atualmente utilizados pela empresa não são os adequados.

Assumindo a Normalidade dos dados do ROSA EVO (tal como acontece hoje) o processo produtivo não é capaz. Através da transformação de Box-Cox ou quando analisado com os indicadores corretos percebe-se que na realidade o estado do processo é muito pior.

Não foi possível concluir qual a razão da Não Normalidade deste PC, pela análise das variáveis programáveis dos ROSA EVO nem pela análise das amostragens de inspeção.

O processo produtivo das estufas ROSA EVO deverá ser estudado e adaptado para que seja possível atingir o cumprimento das especificações. O processo produtivo para se tornar capaz necessita que haja uma redução da variabilidade da humidade percentual das rolhas, tal como a deslocação da média do processo para o valor nominal da especificação.

4.2 Trabalho Futuro

Atualmente nos *softwares* da Qualidade existe uma grande dificuldade na uniformização dos dados mestres, nas várias UI's. Como analisado anteriormente, isto tem implicações a longo prazo com a multiplicação de informação nos históricos que fica refletido no QUALIS. Uma proposta para a resolução deste problema é uma *interface* do *ControLab* com o AS400¹, de modo a que as tabelas com dados mestres sejam automaticamente exportadas para cada *ControLab* das várias UI's. Com isto, as UI's teriam uma linguagem universal, não havendo a necessidade de consolidação dos dados no QUALIS. Também seria evitada a multiplicação de trabalho, pois os dados apenas seriam criados uma única vez no AS400, dado que se propagariam nas oito bases de dados da Qualidade.

¹Como relatado na secção 3.1.1 (página 35), alguns registos deste SI terão de ser revistos. Nomeadamente as *Sales* que possuem uma única designação para várias dependências.

A empresa encontra-se numa fase de avaliação para a implementação dum novo ERP (*Enterprise Resource Planning*) que contemplará os dados das principais Unidades de Negócio, que integrará os dados da Qualidade. No entanto esta proposta deverá ser analisada para responder às presentes necessidades do utilizador.

No *ControLab*, após a adaptação do espaço temporal associado ao ensaio de forças de extração é necessário lidar com a temática da impressão dos relatórios. Caso algum teste tenha a si associado mais que um ensaio, por defeito o impresso é o último ensaio. Pela particularidade da evolução das forças de extração com o decorrer do tempo em garrafa terá de existir a opção de escolha da impressão dos testes realizados nos vários espaços temporais, sendo por defeito impresso o ensaio final para cada espaço temporal.

O *SPC* ainda tem algumas limitações para o tratamento dos dados. Atualmente as suas configurações estão repartidas pelas UI's, não tendo em conta o PC. No momento de definir se as análises serão efetuadas através do desvio-padrão real ou estimado pelos fatores tabelados, o *software* assume o desvio-padrão selecionado para toda a UI em causa. Desta forma, tendo vários PC's na mesma UI todos estarão limitados ao mesmo tipo de cálculo. Esta configuração deverá ser revista para existir a possibilidade de configuração em termos de PC individualizado. Deverá ser possível monitorizar PC's distintos da mesma UI com medidas diferentes, uma vez que estes podem ter capacidades de processo distintas.

Este *software* não dá informação da não Normalidade dos dados presentes nas cartas de controlo, para isso apenas tem a ilustração dos histogramas². No caso de estudo foi notório o enviesamento negativo dos valores individuais, que foram sustentados pelo cálculo dos coeficientes de assimetria e achatamento da distribuição dos valores e até mesmo pelos testes de Normalidade. Não obstante, existem dados que por histograma não é possível aferir a Normalidade destes. Assim, o *SPC* deveria ter implementado mais ferramentas para a validação da Normalidade dos dados em análise.

Atualmente, na empresa todos os cálculos realizados pelo *SPC* são por meio dos valores médios das amostragens de inspeção recolhidas. Este *software* deve permitir que haja a hipótese das análises no CEP serem efetuadas por meio dos valores individuais, uma vez que numa fase de implementação para o controlo inicial é fundamental o conhecimento aprofundado do processo produtivo, que pode ser falseado pelas médias, como relatado neste projeto.

Os PC's das várias UI's deverão ser analisados, para a avaliação da Normalidade dos dados coletados para que haja a utilização de cartas de controlo e índices de capacidade adequados, para que não sejam retiradas conclusões do processo produtivo erradas.

²Para a determinação do número de classes têm em conta a seguinte regra: $\lceil 1 + \log_2 n \rceil = 1 + \frac{10}{3} \log_{10} n$, aconselhado para $n > 30$. No ROSA EVO foram analisados os histogramas tendo com referência a regra de Sturges (mais usual), contudo as conclusões gráficas foram iguais.

Bibliografia

- Ahmad, S., Abdollahian, M., and Zeephongsekul, P. (2008). Process capability estimation for non-normal quality characteristics: A comparison of clements, burr and box-cox methods. In Mercer, G. N. and Roberts, A. J., editors, *Proceedings of the 8th Biennial Engineering Mathematics and Applications Conference, EMAC-2007*, volume 49 of *ANZIAM J.*, pages C642–C665. <http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/357> [June 11, 2008].
- Amorim, C. (2013). Relatório de sustentabilidade 2013 – CORTICEIRA AMORIM, S.G.P.S., S.A. TECHNICAL REPORT, CORTICEIRA AMORIM, MOZELOS.
- APCOR (2011). *Manual Técnico Rolhas*. ASSOCIAÇÃO PORTUGUESA DA CORTIÇA.
- ATHANASOPOULOS, R. J. H. G., RAZBASH, S., SCHMIDT, D., ZHOU, Z., KHAN, Y., AND WANG, C. B. E. (2015). *Package 'forecast'*. VERSION 5.9.
- BONDUELLE, G. (2014). FERRAMENTAS DE CONTROLE. IN *Ferramentas de controle*. DEPARTAMENTO DE ECONOMIA RURAL E EXTENSÃO - DERE, UNIVERSIDADE FEDERAL DO PARANÁ - UFPR.
- BOUMAN, R. AND DONGEN, J. V. (2009). *Pentaho Solutions: Business Intelligence and data warehousing with Pentaho and MySQL*. WILEY PUBLISHING, INC, INDIANAPOLIS.
- BOX, G. E. P. AND COX, D. R. (1964). AN ANALYSIS OF TRANSFORMATION. *Journal of the Royal Statistical Society. Series B (Methodological)*, 46(2):211–252.
- CAVALCANTI, T. R. (2012). SUPORTE A DECISÃO - 02 - SOBRE AS OPERAÇÕES DE OLAP. [HTTP://WWW.ITNERANTE.COM.BR/PROFILES/BLOGS/ARTIGO-SUPORTE-A-DECIS-O-02-SOBRE-AS-OPERA-ES-DE-OLAP](http://www.itnerante.com.br/profiles/blogs/artigo-suporte-a-decis-o-02-sobre-as-opera-es-de-olap). ACEDIDO EM NOVEMBRO 2014.
- CELIEGE (2011). *Código Internacional das Práticas Rolheiras*. CELIEGE EUROPEAN CORK FEDERATION. VERSÃO 6.03.

- CHEN, J.-P. AND DING, C. G. (2001). A NEW PROCESS CAPABILITY INDEX FOR NON-NORMAL DISTRIBUTIONS. *International Journal of Quality & Reliability Management*, 18(7):762–770.
- CLEMENTS, J. A. (1989). PROCESS CAPABILITY CALCULATIONS FOR NON-NORMAL DISTRIBUTIONS. *Quality Progress*, 22(9):95–100.
- COSTA, A. AND RAHIM, M. (2006). A SINGLE EWMA CHART FOR MONITORING PROCESS MEAN AND PROCESS VARIANCE. *Quality Technology & Quantitative Management*, 3(3):295–305.
- CROWDER, S. V. (1987). A SIMPLE METHOD FOR STUDYING RUN-LENGTH DISTRIBUTIONS OF EXPONENTIALLY WEIGHTED MOVING AVERAGE CHARTS. *Technometrics*, 29(4):401–407.
- DA SILVA, D. C. (2011). UMA ARQUITETURA DE BUSINESS INTELLIGENCE PARA PROCESSAMENTO ANALÍTICO BASEADO EM TECNOLOGIAS SEMÂNTICAS EM LINGUAGEM NATURAL. MESTRADO, UNIVERSIDADE FEDERAL DE SANTA CATARINA.
- DILL, S. (2002). UMA METODOLOGIA PARA DESENVOLVIMENTO DA DATA WAREHOUSE E ESTUDO DE CASO. MESTRADO, UNIVERSIDADE FEDERAL DE SANTA CATARINA.
- EGITRON (2014). www.egitron.pt. ACEDIDO EM SETEMBRO 2014.
- ESTEVES, E. (2009). CONTROLO ESTATÍSTICO DA QUALIDADE. IN *Controlo Estatístico da Qualidade*. DEPARTAMENTO DE ENGENHARIA ALIMENTAR, INSTITUTO SUPERIOR DE ENGENHARIA DA UNIVERSIDADE DO ALGARVE, FARO.
- FERNANDEZ, E. S. (2014). *Package 'Johnson'*. VERSION 1.4.
- GEORGE, S. (2012). INMON VS. KIMBALL: WHICH APPROACH IS SUITABLE FOR YOUR DATA WAREHOUSE? [HTTP://SEARCHBUSINESSINTELLIGENCE.TECHTARGET.IN/TIP/INMON-VS-KIMBALL-WHICH-APPROACH-IS-SUITABLE-FOR-YOUR-DATA-WAREHOUSE](http://searchbusinessintelligence.techtarget.in/tip/inmon-vs-kimball-which-approach-is-suitable-for-your-data-warehouse). ACEDIDO EM JANEIRO 2015.
- GOMES, I., FIGUEIREDO, F., AND BARAO, M. I. (2010). *Controlo estatístico da qualidade*. SPE, LISBOA, 2 EDITION.
- GONÇALEZ, P. U. AND WERNER, L. (2009). COMPARAÇÃO DOS ÍNDICES DE CAPACIDADE DO PROCESSO PARA DISTRIBUIÇÕES NÃO NORMAIS. *Gestão & Produção*, 16(1):121–132.

- GORLA, N. (2003). FEATURES TO CONSIDER IN A DATA WAREHOUSING SYSTEM. *COMMUNICATIONS OF THE ACM*, 46(11).
- GOSWAMI, A. AND DUTTA, H. N. (2013). SOME STUDIES ON NORMAL AND NON-NORMAL PROCESS CAPABILITY INDICES. *International Journal of Mathematics and Statistics Invention*, 1(2):31–40.
- HUNTER, J. S. (1986). THE EXPONENTIALLY WEIGHTED MOVING AVERAGE. *Journal of Quality Technology*, 18(4):203–210.
- KHADSE, K. G. AND SHINDE, R. L. (2009). PROBABILITY-BASED PROCESS CAPABILITY INDICES. *Communications in Statistics - Simulation and Computation*, 38(4):884–904.
- KOMSTA, L. (2015). *Package 'moments'*. VERSION 0.14.
- LUCAS, J. M. AND SACCUCCI, M. S. (1990). EXPONENTIALLY WEIGHTED MOVING AVERAGE CONTROL SCHEMES: PROPERTIES AND ENHANCEMENTS. *Technometrics*, 32(1):1–12.
- MONTGOMERY, D. C. (1996). *Introduction to Statistical Quality Control*. WILEY, CANADA, 3 EDITION.
- MORAES, C. F. D., FERREIRA, J. R., AND BALESTRASSI, P. P. (2006). ANÁLISE CRÍTICA DA APLICAÇÃO DE MÉTODOS ESTATÍSTICOS EM PROCESSO DEFINIDOS POR DADOS QUE NÃO APRESENTEM DISTRIBUIÇÃO NORMAL. *GEPROS*, ANO1:7–18.
- OLIVEIRA, C. C. D., GRANATO, D., CARUSO, M. S. F., AND SAKUMA, A. M. (2013). *Manual para elaboração de cartas de controle para monitoramento de processo de medição quantitativos em em laboratórios de ensaio*. SECRETARIA DE SAÚDE.COORDENARIA DE CONTROLE DE DOENÇAS.INSTITUTO ADOLFO LUTZ., SÃO PAULO(ESTADO), 1 EDITION.
- ORACLE (2014). DATA WAREHOUSING CONCEPTS. [HTTP://DOCS.ORACLE.COM/CD/E11882_01/SERVER.112/E25554/CONCEPT.HTM#DWHSG8087](http://docs.oracle.com/cd/E11882_01/server.112/e25554/concept.htm#DWHSG8087). ACEDIDO EM NOVEMBRO 2014.
- OSBORNE, J. W. (2010). IMPROVING YOUR DATA TRANSFORMATIONS: APPLYING THE BOX-COX TRANSFORMATION. *Practical Assesment, Research & Evaluation*, 15.
- PEARN, W. L. AND CHEN, K. S. (1997). CAPABILITY INDICES FOR NON-NORMAL DISTRIBUTIONS WITH AN APPLICATION IN ELECTROLYTIC CAPACITOR MANUFACTURING. *Microelectronics Reliability*, 37(12):1853–1858.

- PORTALACTION (1997–2015). ANÁLISE DE CAPACIDADE. [HTTP://WWW.PORTALACTION.COM.BR/566-%C3%ADNDICES-DE-CAPACIDADE-DO-PROCESSO-CP-E-CPK](http://www.portalaction.com.br/566-%C3%ADNDICES-DE-CAPACIDADE-DO-PROCESSO-CP-E-CPK). ACEDIDO EM JANEIRO 2015.
- RAZALI, N. M. AND WAH, Y. B. (2011). POWER COMPARISONS OF SHAPIRO-WILK, KOLMOGOROV-SMIRNOV, LILLIEFORS AND ANDERSON-DARLING TESTS. *Journal of Statistical Modeling and Analytics*, 2(1):21–33.
- RIBEIRO, J. L. D. AND CATEN, C. S. T. (2012). CONTROLE ESTATÍSTICO DO PROCESSO. IN *Controle Estatístico do Processo*, PORTO ALEGRE, RS. FEENG/UFRGS.
- ROBERTS, S. W. (2000). CONTROL CHART TESTS BASED ON GEOMETRIC MOVING AVERAGES. *Technometrics*, 42:97–101.
- SARKIS, L. (2001). DATA WAREHOUSE: O PROCESSO DE MIGRAÇÃO DE DADOS. MESTRADO, UNIVERSIDADE FEDERAL DE SANTA CATARINA.
- SCRUCCA, L. (2015). *Package 'qcc'*. VERSION 2.6.
- SHAPIRO, S. S. AND WILK, M. B. (1965). AN ANALYSIS OF VARIANCE TEST FOR NORMALITY (COMPLETE SAMPLES). *Biometrika*, 52(3/4):591–611.
- SILVEIRA, C. B. (2013). CARTAS DE CONTROLE. [HTTP://WWW.CITISYSTEMS.COM.BR/CARTAS-DE-CONTROLE](http://www.citisystems.com.br/CARTAS-DE-CONTROLE). ACEDIDO EM DEZEMBRO 2014.
- SPECTRUM, D. (2010). CEP - CONTROLE ESTATÍSTICO DE PROCESSO. [HTTP://WWW.DATALYZER.COM.BR/SITE/SUPORTE/ADMINISTRADOR/INFO/ARQUIVOS/INFO22/22.HTML](http://www.dataalyzer.com.br/site/suporte/administrador/info/arquivos/info22/22.html). ACEDIDO EM JANEIRO 2015.
- THOMSEN, E. (1997). *OLAP Solutions: Building Multidimensional Information Systems*. JOHN WILEY & SONS, INC, NEW YORK.
- TUKEY, J. W. (1957). ON THE COMPARATIVE ANATOMY OF TRANSFORMATIONS. *The Annals of Mathematical Statistics*, 28(3):602–632.
- URBANEK, S. (2012). STAR BROWSER, PART 2: JOINS AND DENORMALIZATION. [HTTP://BLOG.DATABREWERY.ORG/INDEX4.HTML](http://blog.databrewery.org/index4.html). ACEDIDO EM NOVEMBRO 2014.